From THE DEPARTMENT OF PUBLIC HEALTH SCIENCES
Karolinska Institutet, Stockholm, Sweden

EFFECTIVENESS AND COST-EFFECTIVENESS OF EARLY CHILDHOOD OBESITY PREVENTION

Nora Döring

Karolinska Institutet

Stockholm 2018

Published by Karolinska Institutet.
Printed by Eprint AB 2018

ISBN 978-91-7676-960-7

EFFECTIVENESS AND COST-EFFECTIVENESS OF EARLY CHILDHOOD OBESITY PREVENTION

THESIS FOR DOCTORAL DEGREE (Ph.D.)

By

Nora Döring

Public Defence on the 8th of June, 2018 at 9:15
Lecture hall Rockefeller, Nobels väg 11, Karolinska Institutet, Solna Campus

Principal Supervisor:
Professor Finn Rasmussen
Karolinska Institutet
Department of Public Health Sciences

Co-supervisors:
Jeroen de Munter, PhD
Karolinska Institutet
Department of Public Health Sciences

Associate Professor Niklas Zethraeus
Karolinska Institutet
Department of Learning, Informatics, Management and Ethics

Professor Berit L. Heitmann
Copenhagen University
Department of Public Health

Opponent:
Professor Anna Winkvist
Gothenburg University
Department of Internal Medicine and Clinical Nutrition

Examination Board:
Associate Professor Paulina Nowicka
Uppsala University
Department of Food, Nutrition and Dietetics

Associate Professor Anna Månsdotter
Karolinska Institutet
Department of Public Health Sciences

Professor Maria Jansen
Maastricht University
Department of Health Services Research

ABSTRACT

Background: Due to individual health-related and social consequences, but also due to its immense cost burden, obesity remains a large public health challenge for which effective and cost-effective prevention is lacking. It has been suggested that primary preventive efforts operate optimally if started in early childhood, but scientific evidence is scarce. Optimal target groups and intervention programs are most likely to vary by structure of society, population characteristics, health care resources and many other factors.

Aims: The primary aim of this thesis is to evaluative the effectiveness and cost-effectiveness of a primary prevention programme addressing childhood obesity at Swedish child health care centres (CHCs). A secondary aim was to increase understanding of the association between health-related quality of life (HRQoL) and body mass index (BMI), which will be useful for future health economic evaluations.

Methods: Study I provides a description of the PRIMROSE intervention study, which includes a relative validation of a semi-quantitative food frequency questionnaire. Study II is based on repeated self-reported BMI and HRQoL measurements from the Stockholm Public Health Cohort. Study III is a main effect evaluation of the PRIMROSE cluster RCT, which includes 1148 children at follow-up. Study IV is a systematic narrative literature review that explores the methods and applications of economic evaluations within the field of childhood obesity. Study V is a trial-based economic evaluation of the PRIMROSE intervention from a societal perspective.

Results: Compared to normal weight, overweight and obesity was found to be associated with lower HRQoL, especially in the domains of mobility, self-care and usual activity. Heavy weight gain over eight years was also associated with lower HRQoL, whereas weight loss had no protective effect. The PRIMROSE intervention had no significant effect on BMI, waist circumference, obesity prevalence and objectively measured physical activity. Small intervention effects were found in dietary habits but should be interpreted with caution given the possibility of bias in assessments. Study IV showed that there were only a few health economic evaluations of childhood obesity, but these presented already small intervention effect to be worth the money. However, cost-effectiveness was largely dependent on model assumptions and, in particular, on decision-maker's willingness to pay. The economic evaluation presented in Study V showed that PRIMROSE cannot be considered cost-effective given uncertainty around the effect measure.

Conclusion: There was no evidence for the effectiveness or cost-effectiveness of the population-based PRIMROSE RCT. Since the prevalence of childhood obesity remains high, further research is needed to disentangle the "failure" of intervention programs per se from "failure" of implementation and "failure" of evaluation.

LIST OF SCIENTIFIC PAPERS

I. **Döring, N**, de Munter, J, Rasmussen, F. The associations between overweight, weight change and health related quality of life: Longitudinal data from the Stockholm Public Health Cohort 2002–2010. *Preventive Medicine*. 2015;75:12–7

II. **Döring N**, Hansson LM, Scheers Andersson E, et al. Primary prevention of childhood obesity through counselling sessions at Swedish child health centres: design, methods and baseline sample characteristics of the PRIMROSE cluster-randomised trial. *BMC Public Health*. 2014;14(1):335

III. **Döring N**, Ghaderi A, Bohman B, et al. Motivational interviewing to prevent childhood obesity: a cluster RCT. *Pediatrics*. 2016;137(5):1–10

IV. **Döring N**, Mayer S, Rasmussen F, Sonntag D. Economic evaluation of obesity prevention in early childhood: methods, limitations and recommendations. *Int J Environ Res Public Health.* 2016; 911

V. **Döring N**, Zethraeus N, Tynelius P, de Munter J, Sonntag D, Rasmussen F. Economic evaluation of PRIMROSE - A trial-based analysis of an early childhood intervention to prevent obesity. *Frontiers Endocrinology*. 2018; 9:104

CONTENTS

LIST OF ABBREVIATIONS

BMI	Body Mass Index
CBA	Cost-Benefit Analysis
CEA	Cost-Effectiveness Analysis
CHC	Child Health Care Centres
CUA	Cost-Utility Analysis
FFQ	Food Frequency Questionnaire
GEE	Generalised Estimation Equation
HRQoL	Health-related Quality of Life
ICER	Incremental Cost-Effectiveness Ratio
IOTF	International Obesity Task Force
ITT	Intention-To-Treat
MET	Metabolic Equivalent of Task
MI	Motivational Interviewing
MVPA	Moderate to Vigorous Physical Activity
PA	Physical Activity
PP	Per-Protocol
QALY	Quality-adjusted Life Years
RCT	Randomised-controlled Trial
RR	Relative Risk
SCT	Social Cognitive Theory
SD	Standard Deviation
SEK	Swedish Krona
SPHC	Stockholm Public Health Cohort
UK	United Kingdom
USA	United States of America
WHO	World Health Organization

1 INTRODUCTION

Childhood obesity is one of the biggest public health challenges of our time, and primary prevention in early childhood may be the most (cost-) effective approach to it. Yet in practice, little is known about the effectiveness and cost-effectiveness of primary prevention at preschool ages. Indeed, the last decade of empirical research points in a negative direction: primary prevention of early childhood obesity does not seem to work. How can it be that something that *should* work in theory does not work in practice?

In 2008, when the PRIMROSE intervention study commenced, even less was known. Obesity prevalence was at an all-time high, and childhood obesity was a growing concern for health professionals and policy makers alike. At the same time, motivational interviewing (MI) was on the radar of health promotion. Could a technique used for the treatment of addictions also work as a component of obesity prevention?

The primary care setting was suggested as a promising intervention arena for obesity prevention, but was at that time largely underutilised. In Sweden, almost all families attend a child health care centre (CHC) for preventive health care services. However, the focus is mainly on general growth and the provision of vaccination, and only to a lesser extent on young children's developing food habits or physical activity. Since nurses reach families to take up vaccination plans, can we also reach families to follow healthy eating and physical activity behaviour in this setting?

But even if a well-designed intervention became successful in the CHC setting, would it be worth the money? With unlimited needs and scare resources, decision-makers need additional information on how to allocate their money. In order to answer this question, we do not only need data on the costs of implementing an intervention, but we also need to capture all its effects, including societal savings. And how do we do that in the field of early childhood obesity prevention, where the vast majority of positive effects are expected only far into the future?

This dissertation aims to address these issues by presenting the methods and results of a large population-based cluster randomised controlled trial (RCT). The following chapters provide not only summarising but also additional information on background, methodology, and main results in the included papers. The last chapter is a discussion of methods and results of the thesis, which are placed in the context of current research.

2 BACKGROUND

In the following section, I will introduce the main themes of the thesis, including childhood obesity, primary prevention, and decision-making in public health.

2.1 CHILDHOOD OBESITY

2.1.1 Definitions

The World Health Organization (WHO) defines obesity as "abnormal or excessive fat accumulation that may impair health" [1]. Although not as sensitive as some other measurements, the most simple and common way to classify obesity and overweight is the body mass index (BMI), which refers to a person's weight in kilograms, divided by the square of his or her height in meters (kg/m^2). According to an international classification, a BMI greater or equal to 25 but below 30 denotes overweight, while a BMI greater or equal to 30 denotes obesity [2].

These cut-offs are, however, not applicable to children, because children's BMI normally develops over time, with a rapid increase in the first year, followed by a decline up to the age of 5-6 years, and then a gradual increase until adulthood. To take into account development over time, BMI-cut offs for children need to vary by age and gender. The most widely used definition of childhood obesity has been established by the International Obesity Task Force (IOTF) which corresponds to adult cut-off points for overweight and obesity at the age of 18. Yet, cut-offs decrease between 2-4 years of age and then gradually increase from ages 5-6 until age 18, where adult cut-offs are used [3]. An alternative is the WHO standard, which is used for children up to 59 months of age and based on a sample of healthy breastfed children from six countries (Brazil, Ghana, India, Norway, Oman, and the USA). The WHO standard defines overweight as > 2 standard deviations above the WHO growth standard median [4]. A study using a sample of Czech children found that these two different approaches led to different classification, especially among pre-school girls, where, according to the IOTF, 15% were classified as overweight compared to 3% using the WHO standard [5].

2.1.2 Prevalence

Starting in the 1980s, overweight and obesity prevalence rapidly increased; for children at the age of 10, the increase was up to 4-5 times until the early 2000s [6]. During the last decade the prevalence of overweight and obesity in children has stabilised with some indications of a levelling-off [7, 8]. A review from Sweden showed that there was a stabilisation of overweight prevalence among school children 10-11 years of age (born during 1989–1995) [9]. The results from the analysis of the "BEST cohort" supported this observation, showing a stabilisation and decline from 1991 onwards among 8 year-old boys in the Gothenburg region [10]. For children under 5 years of age, there is no systematically collected nationwide data in Sweden. Based on several county-specific reports, however, it seems reasonable to conclude that large variations by region exist [8]. The latest yearly report on children's health from Stockholm County, showed a decrease in overweight prevalence [11]. Among 4-year-old children born in 2011, 8.1% were overweight, compared to 9.4% three years before. The prevalence of obesity in 4-year-old children did not decrease but remained stable at 1.8%. Furthermore, there are large

differences in prevalence of overweight between the municipalities of Stockholm County, ranging from 6% in Lidingö to 16.9% in Skärholmen among 4-year-olds, indicating a strong social gradient in childhood obesity [11].

2.1.3 Modifiable Risk Factors for Childhood Obesity

Physical Activity and Sedentary Behaviour

A lack of physical activity (PA) is known to be associated with several chronic diseases, including obesity. According to the WHO, it is recommended that children (5-17 years) spend at least 60 minutes per day in moderate to vigorous physical activity (MVPA) [12]. For the age group < 5 years, national recommendations only exist in Canada [13], Australia [14] and the UK [15], suggesting at least three hours (180 minutes) of age-appropriate physical activity at any intensity level (i.e., light, moderate, vigorous). In the papers referred to, there are descriptions of what age-appropriate types of activity are. For example, for a toddler, activity is described as "any activity that gets the kids moving", including climbing stairs or playing outside. Yet, intensity and time duration are only described, and not further specified.

As well as lack of physical activity, there has also been an increase in time spent sedentary. Sedentary behaviour is defined as "any waking behaviour characterised by an energy expenditure ≤ 1.5 METs while in a sitting or reclining position" [16], and includes sitting, watching TV, or playing with electronic devices [17]. Over the last decades, there has been a sharp increase in sedentary behaviour in Westernised societies, which is often associated with changing demands in the workplace and increases in screen time [18]. Similar trends have been seen among children. Sedentary behaviour has been associated with a number of negative health outcomes. The question remains whether there are independent effects of sedentary behaviour on health outcomes, as there might be adverse metabolic and health effects of prolonged inactivity even if the recommendation for MVPA is met. A recent systematic review found no evidence for an association between sedentary behaviour and health outcomes in children after adjusting for MVPA [19]. Other studies have, however, shown contrasting results [20, 21]. Furthermore, the behaviours learned in early childhood may continue into adolescence and adulthood [22], where in turn sedentary behaviour is associated with adverse health events even when reaching the recommended level of MVPA [23]. A meta-analysis on the association between physical activity and sedentary behaviour has shown that the two are significantly inversely associated (r=-0.108, 95% CI: -0.128; -0.087). However, this rather weak association indicates that physical activity and sedentary behaviour can coexist, and do not necessarily replace one other, and may therefore need to be targeted separately [24].

Dietary Intake and Eating Habits

When we think of the underlying cause of obesity, i.e., energy intake exceeding energy expenditure, the importance of dietary intake for obesity development is undebatable. Yet, the knowledge and evidence base surrounding energy intake, diet composition, and intake of certain food items, on the one hand, and later overweight in children, on the other, is not fully revealed. This can be partly explained by the methodological challenges measuring eating habits and dietary intake [25]. The only sound evidence for a risk of weight gain and obesity has been seen with sugar-sweetened beverages [26]. High intake of fruits, vegetables, and fish has been associated with a lower risk of a number of diseases [27], but the evidence base for

an isolated effect on childhood obesity remains inconclusive [28]. Children's food preferences and intake patterns are directly and indirectly influenced by their parents [29, 30]. There has been increasing awareness of the effects of eating behaviours beyond the intake of certain food items, e.g., regular family meals. Other parental influences, including pressure by parents to eat (e.g., finishing the plate) or food as a reward have been also associated with food choices [31] and possibly later obesity [32].

Other Factors

There is a growing body of research focusing on identifying early determinants of obesity [33], where the prenatal and the early postnatal periods have been suggested to be "critical periods" for the development of obesity [34]. The term critical period stems from the field of life-course epidemiology, describing limited time periods in which an exposure has an adverse or protective effect on future health outcomes [35]. Among others, these include maternal pre-pregnancy weight [36], maternal socio-economic position, high gestational weight gain [37], gestational diabetes, smoking during pregnancy [38], high birth weight [39], rapid weight gain during the first year of life [40], no or short duration of breastfeeding [41], early introduction of solid foods [42], and short sleep duration [43].

2.1.4 Consequences of Childhood Obesity

Health Consequences

Once people develop overweight or obesity, they are at higher risk of developing chronic disease conditions, such as cardiovascular diseases [44], diabetes [45], musculoskeletal disorders [46], and some types of cancer [47]. Children with overweight and/or obesity not only run an increased risk of becoming obese as adults [48], they are already in childhood likely to develop comorbidities, such as type-2 diabetes [49], hypertension [50], and glucose intolerance [51]. Furthermore, childhood obesity is associated with low self-esteem and can, due to stigmatisation and weight-related teasing, lead to symptoms of depression and (perceived) social rejection [52, 53]. Through the development of these co-morbidities, but even through obesity itself, health-related quality of life (HRQoL) may be significantly impaired [54, 55].

Societal Consequences

As well as health-related consequences, obesity potentially has large societal consequences through direct and indirect costs. In Sweden, direct health care costs related to obesity have been estimated to exceed several billion SEK annually [56]. Additionally, indirect costs, including productivity losses associated with overweight and obesity (sick-leave, disability pension, death before retirement) and other types of exclusion from the labour market (e.g., stigmatisation), are estimated to exceed health-related costs [57-60]. A recent systematic review synthesised the results on lifetime costs due to obesity in childhood and adolescence. The mean total lifetime cost of a child or adolescent with obesity was estimated to be €149 206 (range: €129 410 to €178 933), with the vast majority of the cost being due to productivity losses and, more specifically, income penalties [61]. Five of the 13 included studies were based on adults (≥ 18 years old), and only two studies, from Germany, were based on an early childhood sample [62, 63].

While there have been serious attempts to quantify the economic and societal impact of adulthood obesity in a Swedish setting [60], the economics of early childhood obesity prevention and the long-term economic consequences of childhood obesity in Sweden remain insufficiently explored and evaluated.

2.2 PREVENTION

While there is a fairly good understanding of the physiological cause of obesity (energy imbalance between calories consumed and calories expended), the underlying reasons for this imbalance are multi-dimensional. Knowledge of genetic factors contributing to obesity is increasing. Yet, the almost doubled prevalence of obesity worldwide since the 1980s indicates that environmental factors (e.g., infrastructure, policies) and individual health behaviours (e.g., physical activity, sedentary behaviour, and dietary habits) and interactions between these factors play a vital role in the aetiology of obesity [28]. There is a need to identify feasible, sustainable and cost-effective strategies that can decrease the occurrence of childhood obesity in society. With the (possible) exception of gastric bypass surgery in adults, and possibly also in adolescents [64], there are currently no effective treatment options for obesity, which emphasises the need for prevention.

2.2.1 Stages of Prevention

Traditionally, there has been a differentiation between primary, secondary and tertiary prevention strategies, which can be implemented at both individual and population level. Primary prevention refers to interventions to prevent the occurrence of a disease before it develops [29]. Within primary prevention, some public health researchers further differentiate between primary prevention and primordial prevention, where the latter refers to preventing the emergence of risk factors. Primordial prevention addresses broad health determinants rather than preventing exposure at the individual level [30]. Secondary prevention aims at the early recognition (e.g., by screening) of disease to limit its occurrence, while tertiary prevention focuses on limiting the consequences of disease (i.e., treating existing symptomatic disease) [29]. With regard to the prevention of obesity, the boundaries between the levels of prevention may be less distinct, because overweight/obesity can be considered as both a symptom and a disease. An alternative form of classification has been suggested, differentiating between indicative prevention, selective prevention, and universal prevention [65]. Universal prevention refers to preventive efforts targeting the population at large. Targeting populations subgroups with a high risk set (e.g., defined by ethnicity, socio-economic group or income level) is considered as selective prevention and targeting individuals at high risk as indicative prevention [66]. Instead of regarding these types of classification as distinct, they can be seen as complementary, where the first addresses disease development and the latter the target population.

2.2.2 Primary Prevention of Childhood Obesity

Due to the health consequences of overweight and obesity and the lack of sustainable treatment options, primary prevention seems to be the best way to address childhood obesity in society. Dietary and physical activity habits are established early in life and become less malleable in later life [67]. Parental practices, feeding styles and availability of (un)healthy foods at home,

as well as parents' nutritional knowledge, attitudes and health behaviours, are of major importance for young children's emerging eating and physical-activity habits [68, 69]. Thus, primary preventive efforts are likely to operate optimally if started in early childhood. The involvement of parents in any primary prevention program, as "gatekeepers" for children's access to unhealthy foods and as role models for their emerging dietary habits, is of key importance. It is also important that prevention studies are based on a theoretical framework because it becomes clearer whether and how they are working or not. Yet, interventions that are reported to be based on a theory have not been more effective than interventions developed without a theoretical framework. This may be explained by the fact that theories are not extensively applied in the development of interventions, or are not implemented appropriately [70].

Motivational Interviewing

Motivational interviewing (MI) is an efficient and brief method that prepares people for behaviour change; it is characterised by client-centredness [71]. Eliciting and reinforcing motivation for and commitment to behavioural change are central to MI [71-73]. An MI counsellor's function is not to convince or persuade clients (e.g., parents), but rather to enhance intrinsic motivation, thereby assisting their own decisions about behavioural change. In addition to increasing and maintaining motivation, MI also emphasises the person's self-efficacy, i.e., a person's belief in his/her ability to succeed with a behaviour change. Reflective listening, shared decision-making, and agenda-setting are key components of MI, and counsellors are to be non-judgmental, empathetic, and encouraging [74].

MI appeared in the 1990s, not directly from theory but rather from clinical experience of useful practices in the fields of drug abuse and smoking. Yet, many theoretical influences in the development of MI can be recognised (Figure 1) [75].

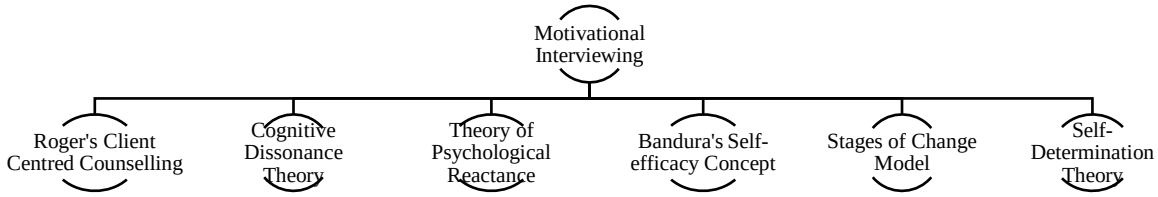


Figure 1 Theoretical Influences Contributing to the Development of MI, based on Söderlund [75]

Social Cognitive Theory

A well-established theory for understanding health behaviour is social-cognitive theory (SCT), which refers to the dynamic interplay between interpersonal factors, behavioural factors, and environmental factors [76]. In intervention research, application of SCT is often characterised by efforts to increase self-efficacy, its key concept [77]. Maternal self-efficacy is a closely related concept defined as a mother's trust in her own capacity to influence her children's health behaviour. It has been shown that maternal self-efficacy is positively associated with children's healthy eating and physical activity [78].

Effectiveness of Obesity Prevention Interventions in Pre-School Children

Due to the difficulties mentioned above, the evidence base for primary prevention of childhood obesity remains weak. Although a Swedish systematic review of the prevention of childhood obesity found some evidence that programs in schools and day-care settings, aimed at promoting healthy eating habits and physical activity, can prevent childhood obesity [79], other systematic reviews have shown no evidence [80-82].

In 2011, a Cochrane review on childhood obesity prevention intervention was published including 55 randomised intervention studies. However, the majority of the included studies were conducted in school settings with children aged 6- to 12-year. Only eight of the studies included children younger than 5 years old, and in seven of these studies children were at least 3 years old at baseline [83]. Although acknowledged as a promising intervention arena, there have been only few attempts targeting preschool children in the primary health care setting [83, 84].

2.3 DECISION-MAKING IN PUBLIC HEALTH

2.3.1 Evidence-based Decision-Making

Randomised Controlled Trials

Randomised controlled trials (RCT) are regarded as the gold standard for assessment of the efficacy of interventions. By randomly assigning individuals to an intervention or control group, the risk of bias is minimised, allowing causal interpretation of the results. While the application of RCT in a clinical setting is fairly straightforward, the use of RCT in a public health setting, especially in the case of interventions based on behavioural change, faces some difficulties. These difficulties are mainly due to the complex causal processes in public health interventions as well as issues concerning ethics and feasibility [43]. Loss to follow-up, a lack of blinding, and cross-over between groups are some of the reasons why internal validity often cannot be assured. This makes it difficult to differentiate between failure of an intervention per se and failure of intervention implementation [44]. And, even in cases where internal validity seems to be good, it remains to be considered whether findings can be implemented in real-life settings.

Cluster Randomised Controlled Trials

Even though the RCT is the most efficient design to study causal relationships, in some instances randomisation at individual level is impractical or inadequate. A cluster RCT is a sub-type of RCT, in which randomisation is not performed at individual level but instead at group (i.e., cluster) level. Cluster RCTs are preferred where the intervention is delivered at an organisational level or unit (e.g., a CHC,) or cannot be directly targeted at individuals (e.g., school class room), or intervention "contamination" is to be expected. One disadvantage of the cluster RCT design is the need for more study participants to reach the same statistical power. Furthermore, due to the correlations within clusters other, more advanced, statistical methods are required to correct for the dependence between observations.

Pragmatic Trials

Pragmatic trials, often embedded in routine practices, "have gained momentum" over the last decades [45]. In contrast to classical (sometimes referred to as explanatory) trials, that measure the efficacy of an intervention, i.e., the effect an intervention generates under ideal circumstances, a pragmatic trial measures effectiveness, i.e., the effect an intervention generates in real-life practice. Pragmatic trials often have a less stringent design which may decrease internal validity. Yet, pragmatic trials are assumed to provide higher external validity in terms of applicability and generalisability. The distinction between these two types of trials is not straightforward. Instead of dichotomising them, pragmatism can be seen as a continuum, where any one prevention study may have a certain degree of pragmatism in its design [85]. The degree of pragmatism may vary according to the research questions to be answered or to contextual factors hindering or facilitating implementation of a more stringent design.

2.3.2 Cost-Effectiveness

A high level of evidence should be the main driver of decision-making in health care. Yet, the costs of a new treatment and prevention programme cannot be neglected in any health care system. In fact, one of the three core principles of the Swedish health care system is the cost-effectiveness principle, which states that the cost of health services should be reasonable from a medical, humanitarian and social-economic perspective [86]. Given the dilemma shared by all publicly financed health care systems, namely scarce resources with unlimited health care needs, prioritisation is necessary. Economic evaluations of existing and new treatment options can facilitate health care decision-making aimed at improving efficiency in the allocation of scarce resources. An economic evaluation may be defined as a "comparative analysis of alternative courses of actions in terms of both their costs and consequences" [87].

Types of Economic Evaluation

It is common to differentiate between four types of economic evaluation: cost-minimisation, cost-effectiveness (CEA), cost-utility (CUA), and cost-benefit analysis (CBA). All four types of evaluation measure costs in similar ways and therefore differ mainly with respect to how (health) effects are measured (see Table 1).

Table 1 Types of economic evaluation, adapted from Drummond et al. [87]

Type of economic evaluation	Valuation of costs	Valuation of effects
Cost-minimisation	Monetary unit	None (alternatives are assumed to have equal outcomes)
Cost-effectiveness	Monetary unit	One-dimensional unit (e.g., BMI)
Cost-utility	Monetary unit	Multi-dimensional unit (e.g., QALYs)
Cost-benefit	Monetary unit	Monetary unit

Cost-effectiveness Analysis and Cost-utility Analysis

A cost-effectiveness analysis (CEA) is one type of full economic evaluation where both the costs and the effects of health care interventions are considered. Both a CEA and a CUA assume that decision-makers strive to maximise gains in a health outcome subject to a cost

constrain [87]. The two types differ only with respect to the assessment of the health consequences. In a CEA the benefits are presented in a one-dimensional unit (e.g., life years gained), whereas in a CUA a generic outcome measure (e.g., quality-adjusted life years) is used, which allows comparison across disease areas [87].

Health-related Quality of Life

Quality-adjusted life years (QALYs) provide a health-outcome measure that is constructed by adjusting life years for the quality of life in which they are spent. This represents an attempt to combine life expectancy and quality of life into a single measure.

The term quality of life spans several interrelated disciplines, with health being one of the core element. The concept of HRQoL takes into account several dimensions of health and has been recognised as an increasingly important, as it goes beyond longevity and also acknowledges the subjective perception of disease and health [18]. Due to complex diseases and chronic disease management instead of the curing of disease, HRQoL becomes a measure that may be more accurate and appropriate in assessing someone's health status. It further allows assessment of health effects before the consequences of symptoms and disease are physically manifested. A year spent in full health refers to 1 QALY (1 year of life x full (1) quality weight=QALY). Following the calculation, half a year spent in perfect years is equivalent to one year spent lived in a health state with a quality weight of 0.5. QALYs are calculated by multiplying life years by a weight representing the HRQoL of health status during those years.

One instrument for measuring HRQoL is the EQ-5D. It has five dimensions, i.e., mobility, self-care, usual activities, pain/discomfort and anxiety/depression with three response levels (1=no problems, 2=some problems, 3=severe problems) defining 243 health states [20]. Each health state can be converted into a single index, applying a so-called value set. Value sets are generated through one or more of the following three techniques: standard gamble, time trade-off or visual analogue scale. While an experience-based value set exists for Sweden [21], there is currently no population-based value set available for the country. The values for the health states in the experience-based value set are assessed from real patients in the health states concerned. For the population-based value set, the general public values hypothetical health states.

Measuring QALYs in Children

While for adults HRQoL has been widely accepted as an endpoint in a research and clinical setting, the assessment of HRQoL among children has only gained awareness and interest in recent years. The assessment of someone's HRQoL relies on his/her subjective evaluation of functioning/impairment in various domains, and it was believed that children's subjective health reports were unreliable and therefore of limited use [88]. However, research indicates that school children (8-10 years of age) [89], and possibly even younger children [90], can adequately reflect and report their health state if instruments are adapted to them. Depending on the children's age, HRQoL measurements may either be self-administered or administered by proxy (e.g., by parents). Due to young children's incompletely developed cognitive and language skills, it is essential to rely on proxy respondents for the assessment of HRQoL in this age group. In a review, Griebsch and colleagues concluded that many aspects of QALY

measurement in children are not fully developed and therefore hinder their application in CUA [91].

2.3.3 Decision-Making beyond Economic Evaluation

As described above, decisions in public health should be based on the best available evidence, and provide the best value for money. But there are other additional factors that decision-makers need to take into account when deciding on the implementation of public health interventions, which may not be directly quantifiable. The ACE-Obesity (Assessing cost-effectiveness in obesity) working group developed so-called second stage filters to evaluate interventions beyond cost-effectiveness. These include the following criteria: "strength of evidence", "equity", "feasibility of implementation", "acceptance by other stakeholders", "sustainability", and "side-effects" [92]. Primarily developed for the assessment of childhood obesity interventions, the criteria are also applicable to decision-making regarding other public health interventions.

3 AIMS

3.1 OVERALL GOAL

The overall objective of this thesis is to evaluate the effectiveness and cost-effectiveness of a large primary prevention programme addressing childhood obesity.

3.2 SPECIFIC OBJECTIVES

Specific objectives are to:

1. Investigate the shape and strength of associations between overweight status, weight change and health-related quality of life.

2. Describe the design and methodology of the PRIMROSE cluster-randomised controlled trial.

3. Validate a semi-quantitative FFQ against a 7-day food diary.

4. Evaluate the effect of the PRIMROSE intervention on 4-year-old children's and their mothers' BMI, dietary pattern and physical activity.

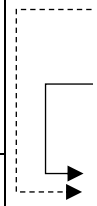
5. Explore and review existing methods and applications of economic evaluations within the field of early childhood obesity prevention.

6. Estimate the costs associated with implementation of the PRIMROSE intervention programme and conduct a trial-based cost-effectiveness analysis.

4 METHODS

4.1 OVERVIEW

Table 2 Overview of the studies included in this dissertation

Evaluation Phase		Study Design	Materials	Studies
Pre-Trial		Study Protocol	PRIMROSE n=1039	II
		Validation Study	n=214	II
Trial		Cluster RCT	PRIMROSE n=1148	III
Supplementary		Cohort Study[1]	Stockholm Public Health Cohort n=16666	I
		Systematic Literature Review	Peer-reviewed articles n=6	IV
Post-Trial		Cost-Effectiveness Analysis	PRIMROSE n=1148	V

[1] Study I was initially intended to inform Study V. However, instead of a cost-utility analysis, we conducted a trial-based cost-effectiveness analysis.

4.2 STUDY I

4.2.1 Study Design

Study I is a prospective cohort study with repeated measurements of BMI and HRQoL. The study was designed to investigate the associations between overweight, weight gain and HRQoL.

4.2.2 Materials

Stockholm Public Health Cohort

The Stockholm Public Health Cohort (SPHC) provided the basis for a prospective study set up within the framework of the Stockholm Public Health Surveys coordinated by the Centre for Epidemiology and Community Health, Stockholm County Council. The cohort was initiated in 2002 and followed up twice (2007, 2010). The total population within Stockholm County of approximately 2 million makes it the largest and most densely populated county in Sweden. The cohort consists of individuals 18–84 years of age, who were selected using a stratified random sampling design from the total population of Stockholm County, with stratification for residential municipality. In 2002, 49 909 individuals were invited to respond to a questionnaire covering HRQoL, height, weight, socio-economic position, chronic diseases and disabilities. 31 182 individuals took part in the survey in 2002. Of those, 23 794 (79%) also responded to the questionnaire in 2007, and 19 128 (61%) in 2010.

Record Linkage

Using the unique personal identification number assigned to all residents in Sweden, the self-reported data in the SPHC has been linked to several registers containing demographic and socioeconomic information, an also data from the National Patient Register and the Swedish Cancer Register. Since 1987, the National Patient Register has covered almost all inpatient care and since 2001 almost all specialised outpatient care in Sweden. Its quality was considered good in a recent evaluation, which found less than 1% underreporting [93]. Primary care, however, is not yet included in the Patient Register. Since 1958, the Swedish Cancer Register has information on all cancer cases in the Swedish population. The reporting of diagnosed cancer cases is compulsory. Underreporting is estimated to be less than 4% [94].

4.2.3 Data Processing and Analysis

For the first research question in Study I, on the overall relationship between overweight status and HRQoL, we pooled observations from 2002 and 2010. Overweight status was determined by BMI, which was calculated from self-reported weight and height, and categorised according to the international classification into underweight, normal weight, overweight and obese. HRQoL was measured using the descriptive system of the EQ-5D instrument, which has five dimensions: mobility, self-care, usual activities, pain/discomfort and anxiety/depression with three response levels (1="no problems", 2="some problems", 3="severe problems"). A weighted summary index was calculated by using a population-based value set. In the absence of an appropriate Swedish population-based value set, the UK value set was used, as recommended [22].

The following variables were considered as potential confounding variables and included in the models: smoking, mental health, chronic disease, civil status, highest attained education and occupational status. The cross-sectional association was analysed using generalised linear models with a Poisson distribution. Poisson distributions are usually applied to count data, but they can also be used for binary outcomes with a log link and robust variance to retrieve a direct estimate of relative risk instead of an odds ratio. This allows for easier interpretation [95]. To model the association between overweight status and the EQ-5D index as a continuous outcome variable, we performed linear regression with a generalised estimation equation (GEE) and adjusted standard errors.

Table 3 Weight change categories between 2002 and 2010

Category	Cut-off
Heavy weight decrease	≤ -10%
Moderate weight decrease	> -10% – < -5%
Stable weight	≥-5% – 5%
Moderate weight increase	≥ 5% – < 10%
Heavy weight increase	≥ 10%

For the second research question on the longitudinal association between weight change and HRQoL, we established weight-change categories, based on percentage weight change relative to baseline weight (Table 3). The same statistical analyses as described above were conducted. In addition to the above-mentioned covariates, an interaction term (weight change in

percentage * baseline BMI category) was included to analyse whether the effect of weight change differs with respect to baseline BMI category.

4.3 STUDY II

4.3.1 Study Design

Study II is a protocol paper, which describes the design and methods of the PRIMROSE cluster RCT (Study III). This entails a (relative) validation study of a semi-quantitative FFQ.

4.3.2 Materials

PRIMROSE

The design and methods of the PRIMROSE trial are described in detail in Study II. There follows a summary of its main aspects.

PRIMROSE is a pragmatic cluster RCT, which was conducted at CHCs in eight counties (Stockholm, Uppsala, Södermanland, Örebro, Gävleborg, Västernorrland, Västmanland, and Jämtland) of Sweden. The trial commenced in 2008 and collection of follow-up data was completed in 2015. 1:1 Randomisation was performed at CHC-unit level, after nurses' written consent to participate. During the recruitment period, there were about n=2 230 families with first-born children receiving care at the participating CHCs. Of those, 1 867 families were eligible. Families were not asked to participate if they did not speak Swedish, were about to move/change CHC, or had severe family situations making it unethical to ask them to participate in the trial. When children were 5–6 months of age, trial nurses approached eligible families, and n=1 369 agreed to participate.

The intervention was based on SCT and involved 8 MI sessions led by trained nurses aiming to promote healthy eating and PA behaviours. During these sessions, parents, together with the nurses, formulated goals with regard to PA and healthy eating habits, which were then discussed in the following sessions according to the PRIMROSE manual (Figure 2). Families belonging to the control CHC units were offered regular routine health check-ups only. Of the 601 families assigned to the intervention, 543 attended at least one session, and 388 took part in all nine sessions.

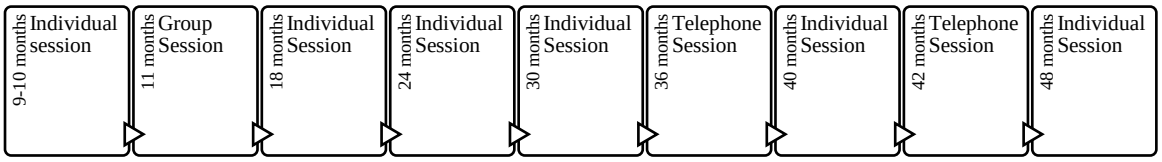


Figure 2 Overview of the PRIMROSE Intervention

Validation Study

To validate the FFQ developed for the PRIMROSE intervention, a separate study was conducted of a random population-based sample of families living in the same counties as the PRIMROSE population. Of the 2 400 invited families, 514 agreed to participate (21.5%). Of those, 68% completed the first food diary. The second food diary, which was sent to those who had filled in the first food diary, had a response rate of 78%. The FFQ was then sent to those families that completed at least one food diary. In total, 288 participants recorded their food intake for at least four days and answered the FFQ. The analytical sample for the validation study was restricted to mother-and-child pairs that completed both food diaries (8 days) and the FFQ (n=214).

Table 4 Descriptive variables of children and mothers included in the validation study of the FFQ

	N	Mean	sd	Min	Max
Child age (years)	214	4.0	1.0	2.8	5.5
Child BMI (kg/m^2)	195	15.8	1.6	12.4	23.0
Maternal age (years)	214	37.2	4.4	24.6	47.7
Maternal BMI (kg/m^2)	212	23.3	3.8	17.1	41.2

4.3.3 Data Processing and Analysis

To compare baseline characteristics between the intervention and control groups, t-tests for continuous variables and $\chi 2$-tests for categorical variables were performed. These were derived from unadjusted linear and multinomial regression models, with a robust (sandwich) estimator to take clustering into account. Standard statistical techniques are inappropriate for repeated or clustered data, since a key assumption (i.e., the assumption of homoscedasticity) is violated in a clustered RCT. By using the approach outlined above, we were able to estimate valid standard errors without requiring the usual assumption of the residual errors to have constant variance. Analyses of baseline characteristics can indicate whether randomisation has been successful, or whether there might be some selection bias. However, there is a debate over the statistical testing of baseline randomisation group differences [96]. In fact, according to the CONSORT statement, it is recommended that significance testing of baseline differences is not performed, since it may "mislead investigators and their readers" [97].

For the relative validation of the FFQ, mean and median frequencies of food intakes were calculated for both the FFQ and the food diaries. To analyse the agreement of the two methods of food intake assessment, Spearman rank-correlation coefficients were calculated, and interpreted according to widely accepted rule-of-thumb cut-offs [98] (Table 5).

Table 5 Interpretation of strength of correlation [98]

Very weak	**Weak**	**Moderate**	**Strong**	**Very strong**
0.00-0.19	0.20-0.39	0.40-0.59	0.60-0.79	0.80-1.0

4.4 STUDY III

4.4.1 Study Design

Study III is concerned with main effects in the cluster RCT PRIMROSE. The design and methods of the PRIMROSE trial are described in detail in Study II, and the following will therefore focus on the data processing and analysis only.

4.4.2 Data Processing and Analysis

The primary outcome of PRIMROSE was children's BMI at age 4. Measurements were taken by study nurses at all visits to the CHCs, but BMI was not always exactly measured at age 4. Therefore, we applied growth curve modelling, using non-parametric regression (kernel smoothing) to estimate children's BMI at age 4 [99]. Secondary outcomes were mother's BMI, children's and mother's food intake and PA.

To evaluate the intervention's effectiveness, both intention-to-treat and per-protocol analyses were performed. We did not perform any imputation of missing variables, and therefore did the ITT analysis only to the extent that missing values allowed (i.e., having at least the primary outcome measured at follow-up). In the per-protocol analysis, we restricted the sample to the families that completed the entire intervention, delivered by nurses who had completed their MI training.

We ran unadjusted linear regressions on continuous variables and log binominal regressions on binary variables using generalised estimating equations with robust variance estimates, taking into account the cluster-randomised study design.

4.5 STUDY IV

4.5.1 Study Design

Study IV describes a systematic narrative literature review designed to explore existing methods and applications of cost-effectiveness analyses in the field of early childhood obesity prevention.

4.5.2 Materials

The systematic review was conducted in line with the guidelines of the Centre for Reviews and Dissemination [100]. We searched for relevant articles published between 2004 and November 2015 using the main electronic databases for health sciences and electronic evaluations (PubMed, Cochrane Library, NHS Economic Evaluation Database, and EconLit). Based on our research question, and key terms and articles identified beforehand, an independent librarian developed and applied a search strategy which led to n=728 articles. Based on title and abstract screening, 717 articles were excluded, leaving 11 articles for full-text screening. The final sample consisted of 6 studies.

Table 6 Inclusion and Exclusion Criteria

Inclusion	Exclusion
Trial-based cost-effectiveness analyses, simulation/model-based cost-effectiveness analyses	Reviews, meta analyses, qualitative studies, partial economic evaluations (i.e., cost descriptions), not peer reviewed
Target population: Preschool children (<6 years) and/or their parents	Selected target groups (e.g., low SES, ethnic groups)
Intervention: Behavioural intervention targeting dietary and physical-activity behaviours	Pharmaceutical intervention, surgical intervention, structural intervention
Intervention outcome measures must include at least one of the following: BMI or waist circumference, overweight prevalence	
Language: English or German	No abstract available
European countries, USA, Canada, Australia, New Zealand	Developing countries

4.6 STUDY V

4.6.1 Study Design

Study V is a trial-based economic evaluation of the PRIMROSE cluster RCT (Study III) performed from a societal perspective.

4.6.2 Materials

Study V is based on data from the PRIMROSE cluster RCT, which has been described earlier (see section 4.3). For Study V, we restricted the sample to the families that had completed the interventions (i.e., with available follow-up measurements). Average costs for the teaching and training of nurses and costs of intervention delivery were included.

Table 7 Duration of visits and attendance among the intervention group that participated in at least one session (n=543)

	Visit 1	Visit 2	Visit 3	Visit 4	Visit 5	Visit 6	Visit 7	Visit 8	Visit 9
Type	I	G	I	I	I	T	I	T	I
Uptake	100%	93%	90%	86%	83%	75%	75%	65%	70%
Length	65 min	84 min	52 min	53 min	51 min	24 min	51 min	22 min	50 min
Visitor									
Mother	51%	56%	51%	53%	60%	86%	63%	84%	62%
Father	3%	8%	10%	12%	5%	9%	10%	14%	11%
Both	46%	36%	39%	36%	35%	5%	28%	3%	27%

I=individual session, G=group session, T=telephone session

4.6.3 Data Processing and Analysis

Costs and effects were estimated from participant-level data. The effect measure for Study V was children's BMI at age 4, and based on the calculation described above in the section about Study III (section 4.4). The difference between groups was interpreted as BMI prevented.

Costs were calculated for both the intervention and the control group. For the intervention group, costs included costs for education, training costs, and implementation costs. Based on invoices, training costs were calculated for one 5-day workshop, and then assumed to be similar for the other workshops. In order to estimate the costs for intervention delivery, data on the attendance and length of the individual sessions were recorded by the participating nurses (Table 7). This allowed us to estimate implementation costs for each participating family separately. Unfortunately, we did not collect any data on employment status and/or salary. Instead, we estimated salaries by using the age-specific average salaries for the regions provided by Statistics Sweden. Costs for the control group cover the regular check-up meetings only, and were estimated in the same way as in the intervention group.

Cost-effectiveness was expressed by the Incremental Cost-Effeteness Ratio (ICER), which is calculated as follows:

$$ICER = \frac{(C_1 - C_2)}{(E_1 - E_2)}$$

C_1 = Costs in intervention group
C_2 = Costs in control group
E_1 = Effects in intervention group
E_2 = Effects in control group

To indicate the level of uncertainty around the estimate, non-parametric bootstrapping was applied, where 1000 cost and outcome pairs were generated (with replacement) [18]. These results were then displayed on a so-called cost-effectiveness plane. A cost-effectiveness plane allows graphical representation of results and the uncertainty surrounding the estimate. Further, it allows a better interpretation of the ICER, since a negative value of the ICER can be both positive (lower costs with higher effects) and negative (higher costs with lower effects). The cost-effectiveness plane consists of a four-quadrant diagram, where the x-axis represents incremental effects, and the y-axis incremental costs. For quadrants IV and II, the choice between the interventions is clear. When intervention I is more effective and costs less than intervention II (quadrant II), it can be said that alternative I dominates alternative II. However, the vast majority of cost-effectiveness analyses present results in quadrant I, which means that a new intervention generates more health gains but at higher costs. In these situations, the choice is less clear and depends on the decision-maker and the willingness-to-pay (WTP). WTP is the maximum price a decision-maker is prepared (i.e., willing) to pay for a given outcome, and therefore regarded as a threshold for cost-effectiveness.

To shed light on the probability of cost-effectiveness, the results were then illustrated by using cost-effectiveness acceptability curves (CEACs), in which the probability that the PRIMROSE intervention is cost-effective was displayed for different theoretical WTP levels for the prevention of 1 BMI unit.

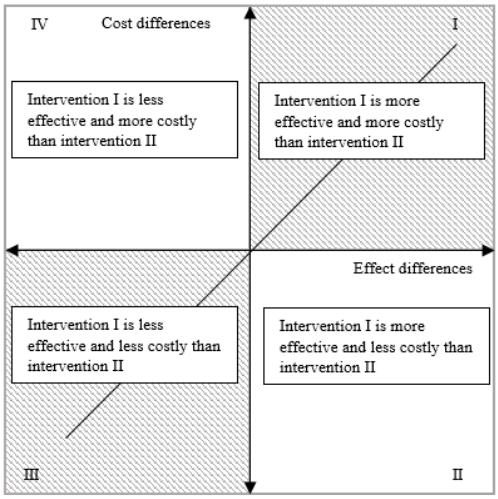


Figure 3 Cost-effectiveness plane, based on Drummond [32]

4.7 ETHICAL CONSIDERATIONS

Ethical Approval

All studies included in this thesis have been approved by the Ethical Review Board in Stockholm, Sweden.

Ethical Concerns with the SPHC

For Study I which is based on data from the Stockholm Public Health Cohort, we do not see any substantial ethical issues. The participants in the SPHC consisted of a random sample of individuals aged 18–84 years living in Stockholm County. Invitations to participate were sampled from the Register of the Total Population, stratified by residential municipality. Participants gave their informed consent to the use of their national registration numbers and future contacts, and record linkages.

Ethical Concerns with the PRIMROSE Intervention Study

Parents provided written informed consent and were informed that they could leave the study at any point in time without giving a reason. Since the children were only 6 months of age when families were approached by our study team, parents also consented on behalf of their children.

Although it seems very unlikely that taking part in the study has had adverse effects on participating parents and children, one possible adverse effect concerns additional stress for participating families. The first months/years after the birth of a child are a vulnerable time for parents. To avoid an additional burden for first-time parents, only families without severe social difficulties or severe health constraints were approached for dialogue about participation. However, even families without any social/health constraints can feel insecure, and by creating

awareness of a potential harmful health behaviour, it cannot be ruled out that parents might feel ashamed, or even more insecure, about their own behaviour or parental skills. During the intervention period, parents were offered constant reflection over their own behaviour and guidance on how to promote healthy behaviour. Therefore, we believe that the benefits of participation would outweigh (or prevent) the potential negative effects of participation.

5 RESULTS

5.1 WEIGHT CHANGE AND HEALTH-RELATED QUALITY OF LIFE

Adjusted pooled linear regression analysis showed that people being overweight or obese, respectively, had -0.015 (95% CI: -0.018; -0.010) and -0.039 (95% CI: -0.047; -0.031) lower EQ-5D indexes compared to those of normal weight. When looking at the specific domains, there were significant elevated risks of reporting problems in all but one domain (i.e., the anxiety/depression domain) among the overweight or obese compared to those of normal weight.

From baseline (year 2002) to follow-up (year 2010), 54.7% of the cohort remained stable in weight (+/- 5%) while more than 30% gained more than 5% of their baseline body weight. Weight gain was significantly associated with lower self-reported HRQoL, which was visible in both in the respective domains and the overall weighted summary score and irrespective of the baseline weight category. Weight loss was not significantly protective for reporting impairments. In contrast, those who lost weight and were normal weight at baseline were up to three times more likely to report problems in the respective HRQoL domains compared to those who remained stable in their normal weight. Only among those who were obese at baseline, there was an indication that moderate weight loss can be protective. However, these effect sizes did not reach statistical significance.

5.2 BASELINE CHARACTISTICS OF THE PRIMROSE STUDY POPULATION

Of the 1 867 families eligible for participation, 1 355 (with 1 369 children) agreed to participate. After retrieval of informed consent and disclosure of group allocation, families received the baseline questionnaire prior to their first intervention meeting at the participating CHCs. Due to time delays, not all (n=489 in the intervention group and n=550 in the control group) mothers filled in the baseline questionnaire prior to the first (intervention) meeting. The baseline data of those who submitted the baseline questionnaire late are not considered in the present analysis (n=46 in the intervention group and n=87 in the control group). In total, the baseline sample consisted of 1 039 mothers with 1 053 children. When comparing the groups, we did not find any statistically significant group differences with the exception of country of birth. Even though not statistically significant, the proportion of women in the intervention group with a university education was 7% larger than in the control group, which may indicate some selection bias. With regard to diet and physical activity, mothers in the two groups were similar. Only for sugared drinks and snacks (discretionary calories) did participating mothers in the intervention group have a lower, although not statistically significant, reported food intake compared to mothers in the control group.

5.3 VALIDATION OF THE FOOD FREQUENCY QUESTIONAIRE

Spearman correlation coefficients varied from weak (e.g., fish and French fries intake) to strong (e.g., mother's fruit intake) (Table 8). In general, there was a systematic increase in differences (both positive and negative) between the two methods with increasing intake. Furthermore, mothers with normal weight, compared to mothers with overweight, tended to be more accurate in the reporting of their own and their children's intake.

Table 8 Relative Validation of a Food Frequency Questionnaire

	Food diary and FFQ				
	n	d†	95% CI	r	95% CI
Children					
Fruits (t/d)	210	0.30	0.19; 0.42	0.42	0.31; 0.53
Vegetables (t/d)	212	1.08	0.96; 1.19	0.48	0.36; 0.58
Fish total (t/w)	211	0.54	0.33; 0.74	0.31	0.18; 0.42
French fries (t/m)	210	0.14	-0.12; 0.40	0.36	0.24; 0.48
Sugared drinks* (t/w)	211	-0.39	-0.65; -0.13	0.59	0.50; 0.67
Discretionary calories** (t/w)	212	0.75	0.18; 1.33	0.56	0.46; 0.65
Mothers					
Fruits (t/d)	213	0.48	0.37; 0.59	0.60	0.50; 0.68
Vegetables (t/d)	213	1.44	1.29; 1.59	0.32	0.19; 0.43
Fish total (t/w)	205	0.36	0.13; 0.60	0.41	0.28; 0.51
French fries (t/m)	211	0.23	0.00; 0.46	0.34	0.22; 0.46
Sugared drinks* (t/w)	209	-0.13	-0.37; 0.11	0.43	0.32; 0.54
Discretionary calories** (t/w)	198	0.19	-0.45; 0.83	0.52	0.41; 0.61

r=Spearman correlation, t=times, d=day, w=week, m=month, CI=confidence interval, *sugared drinks include soda with sugar or lemonade, and chocolate drinks); **discretionary calories include savoury snacks, sugared drinks, sweets, chocolate, pastries, cake and ice cream; † Difference between mean (FFQ) and mean (FD).

5.4 MAIN EFFECTS OF THE PRIMROSE INTERVENTION

At follow-up at age 4, there were 1 148 children with data on weight and height. Mothers who were pregnant at follow-up were excluded (n=82). There were no statistically significant group differences in BMI (mean difference: -0.11, 95% CI: -0.31; 0.08) or waist circumference (mean difference: -0.48, 95% CI: -0.99; 0.04) in the intention-to-treat analysis. In the per-protocol analysis, where the analytical sample was restricted to families who had completed all sessions with nurses (and who had also completed their full training programme (n=1 088)), the results for BMI and overweight prevalence remained virtually the same, whereas the results for WC reached statistical significance.

For the secondary outcomes, there were small, yet significant intervention effects with regard to the intake of some of the food items, i.e., vegetables, French fries, sugared drinks and discretionary calories (Table 10). There was no evidence of an intervention effect with regard to physical activity and sedentary behaviour (data shown in Study III). A similar pattern was observed for mothers. There were also no intervention effect on mother's BMI, WC, or PA (data shown in Study III).

Table 9 Primary Outcome of the Primrose Intervention: Results from the Intention-to-treat and Per-Protocol analyses

	Intervention (n=448)	**Control (n=700)**		
Intention-to-treat	Mean (se)	Mean (se)	Mean Δ/RR (95% CI)	p-value
Children				
BMI (kg/m^2)	16.0 (0.08)	16.1 (0.06)	-0.11 (-0.31; 0.08)	0.26
Waist circumference (cm) [a]	52.5 (0.21)	53.0 (0.26)	-0.48 (-0.99; 0.04)	0.07
Overweight [b] (%)	14.8 (0.02)	15.5 (0.01)	0.95 (0.69; 1.32)	0.78
	Intervention (n=388)	**Control (n=700)**		
Per-Protocol	Mean (se)	Mean (se)	Mean Δ /RR (95% CI)	p-value
Children				
BMI (kg/m^2)	15.9 (0.10)	16.1 (0.08)	-0.15 (-0.35; 0.54)	0.15
Waist circumference (cm) [c]	52.4 (0.24)	53.0 (0.16)	-0.63 (-1.19; -0.70)	0.03
Overweight [b] (%)	14.5 (0.02)	15.5 (0.01)	0.93 (0.66; 1.64)	0.70

[a] missing: n=47, [b] including obesity, [c] missing: n=20

Table 10 Intention to treat analysis of food habits at follow-up of children and mothers in the PRIMROSE trial

	Intervention (n=412)	**Control (n=595)**		
	Mean (se)	Mean (se)	Mean Δ (95% CI)	p-value
Children				
Fruits (t/d)[a]	1.1 (0.03)	1.1 (0.03)	0.01 (-0.09; 0.11)	0.78
Vegetables (t/d)[a]	1.0 (0.03)	0.9 (0.03)	0.13 (0.04; 0.22)	0.01
Fish (t/w)[a]	1.6 (0.06)	1.5 (0.06)	0.10 (-0.06; 0.27)	0.21
French fries (t/m)[a]	1.5 (0.07)	1.8 (0.07)	-0.37 (-0.58; -0.17)	<0.001
Sugared drinks (t/w)[b]	2.2 (0.18)	2. 7 (0.15)	-0.49 (-0.97; -0.15)	0.04
Discretionary calories (t/w)[b]	5.3 (0.17)	5.9 (0.12)	-0.60 (-0.14; -0.25)	0.01

t=times, d=day, w=week, m=month. Sugared drinks include soda with sugar or lemonade, and chocolate drinks; discretionary calories include savoury snacks, sugared drinks, sweets, chocolate, pastries, cake, and ice cream. [a] missing: n=5, [b] missing: n=14

5.5 A SYSTEMATIC REVIEW OF ECONOMIC EVALUATIONS OF EARLY CHILDHOOD OBESITY PREVENTION INTERVENTION

Out of 728 studies that were identified, only six were in line with our inclusion and exclusion criteria. Even though one study did not report any results, it did present the methods of a full economic evaluation and was therefore included. None of the economic evaluations were based on interventions that reported significant effect sizes. Yet, three of them still reported cost-effectiveness (Table 11). Three main conceptual and methodological limitations of economic evaluations were identified: an insufficient conceptual approach considering the complexity of childhood obesity, inadequate measurement of effects of interventions, and a lack of valid instruments to measure the child-related quality of life and costs.

Table 11 Characteristics of the studies included

Authors	**Type**	**Study perspective**	**Time horizon**	**Effectiveness** Mean Δ (95% CI)	**Cost-effectiveness**
Moodie et al.	CEA	Societal	Lifetime	-0.25 (-0.62; 0.12)	Yes
Wake et al.	CCA	Reported: Health care Assumed: Societal	Short (1 year)	-0.00 (-0.50; 0.50)	No
Ma and Frick		Not reported Assumed: Health care	Lifetime	n.a.	n.a.
Moodie et al.	CEA	Societal	Lifetime	-0.28 (-0.70; 0.15)	Yes
Hayes et al.	CEA	Health care	Short (2 years)	0.33 (-0.04; 0.66)	Yes
Pil et al.	CEA/ Design	Societal	Lifetime	n.a.	n.a.

CEA=cost-effectiveness analysis, CCA=cost consequences analysis, n.a= not applicable

5.6 ECONOMIC EVALUATION OF THE PRIMROSE INTERVENTION

The estimated mean total costs per participant in the intervention group were 453 Euros (Min=177, Max=740) in comparison to 111 Euros (Min=0, Max=246) in the control group for usual care. The mean additional costs for carrying out the intervention programme were 342 Euros (95% CI: 334; 348) per participant. The main costs were costs of the education, costs of MI training and supervision, and costs of implementation of the intervention program. The largest component of the PRIMROSE costs arose from delivery of the intervention within the CHC settings. The point estimate of the incremental cost effectiveness ratio (ICER) was 3 109 Euros per 1 BMI unit prevented. About 11% of the bootstrapped pairs were dominated, meaning that the PRIMROSE intervention costs more for less effect. Yet, the vast majority of the bootstrapped ICER estimates indicate increased benefits and greater costs. Given the uncertainty around the effect measure combined with considerable opportunity costs, the current trial-based economic evaluation of PRIMROSE suggests that resources might be better used elsewhere within the field of obesity prevention.

6 DISCUSSION

In what follows, I summarise the main findings of our research, discuss limitations and the findings in light of current discourse and ways ahead, with regard to both childhood obesity prevention in general and more specifically the analysis and (economic) evaluation of intervention studies within the field.

6.1 MAIN FINDINGS

1. Weight gain was significantly associated with lower HRQoL (Study I).

2. There was no preventive effect of weight loss on HRQoL, irrespective of baseline weight (Study I).

3. The FFQ showed moderate to good validity for the majority of food items (Study II).

4. There was no significant intervention effect on children's BMI at age 4 (Study III).

5. There were small significant intervention effects on parent-reported intake of certain food items, but no effects on objectively measured physical activity (Study III).

6. Despite a large and increasing number of intervention studies in early childhood, few have evaluated economic aspects (Study IV).

7. Even small intervention effects can give value for money, but this depends largely on model assumptions and the decision-makers' willingness to pay (Study IV/V).

8. Given the uncertainty around the effect measure, the PRIMROSE intervention programme cannot be considered as cost-effective (Study V).

6.2 LIMITATIONS

To be able to discuss the results from the studies, there is a need to look carefully at methodological issues and limitations.

6.2.1 Study I

In contrast to the RCT design that takes care of confounding by design, observational studies need to take confounding into account when analysing and interpreting the associations of interest. If there is information on all possible confounding variables and a sample large enough to include all covariates to make reasonably precise estimations, the causal impact can be estimated. Yet, in reality, this is often not the case. In the association between overweight and HRQoL, there are a number of factors that may explain (parts of) the observed association. Some may be on the causal pathway, mediating the association of interest. Others may moderate the association, meaning that the effect can be amplified or weakened. And again, some factors may confound the association, i.e., influence both the dependent and independent variable thereby giving a spurious association. A plausible confounder of the association between overweight status (or weight change) and HRQoL is (chronic) disease status. Certain

chronic diseases (e.g., cancer) are known to decrease weight at late stages. In addition, there are other diseases, e.g., psychiatric disorders that are treated with psycho-pharmacological drugs, which result in weight gain and thus overweight. At the same time, having a chronic disease will most likely influence a person's subjective perception of HRQoL. By lacking information about chronic disease or failing to take such information into account, associations may be over- or under-estimated. In the current study, we tried to address this issue by first excluding those with self-reported chronic disease at baseline and then by including self-reporting chronic disease as a covariate in the statistical analyses. However, we could not fully differentiate between intended and unintended weight change.

Another common problem with cohort studies is selection bias due to differential non-response or drop-out. A majority of studies indicate that non-responders have higher morbidity and mortality than those responding to surveys. Indeed, in our sample, we saw that people who were lost to follow-up reported a lower HRQoL at baseline. Differential non-response bias regarding questions on weight and height may also occur. In our analysis, we excluded people who had missing information on the questions of weight and/or height. Again, the response pattern may be determined by weight status, meaning that people with overweight or obesity are less likely to respond to weight-related questions than those of normal weight. In addition to differential non-response bias, our analyses may be further biased by differential misclassification. People with overweight and obesity are more likely to underestimate their weight and overestimate their height in comparison to normal-weight individuals [101].

6.2.2 Study II

Validation studies are intended to provide information on how well an instrument measures what it is supposed to measure. Ideally, in validation studies, we compare our instrument of interest to a so-called gold standard, i.e., the instrument that provides a true measure of condition or phenomenon of interest. For the assessment of dietary intake or eating patterns there is, however, no true gold standard available, and almost all traditional methods rely on self-reported (or proxy-reported) information. The most common are 24-hours food recall, food frequency questionnaires (measuring types and amounts of food, and also intake frequency) and food diaries. Yet, food diaries can be burdensome and scarcely possible to maintain with large samples of individuals. Furthermore, research has indicated that the actual process of food recording may even alter people's food intake. A prominent other method, especially in large cohort studies, is the FFQ, often in semi-quantitative form. It includes questions on a battery of food and beverage items regarding intake frequency and portion sizes. Due to the above-mentioned lack of a true gold standard, validation studies in a strict sense cannot be conducted. Instead, the term relative validation (or validation) is used to describe a validation using imperfect reference instruments, comparative which is commonly performed in the field of nutritional epidemiology. Traditionally, food records have been believed to be more accurate, and the 7-day food diary is often used as the gold standard for validating other methods. Given the inherent limitation of this method, there is a need for care in interpreting the results. For example, our research indicates a good relative validity of fruit intake. This may indicate that our semi-quantitative FFQ accurately measures fruit intake. However, it may indicate alternatively that both instruments provide biased estimates to the same extent, and/or in the same direction, when assessing fruit intake.

6.2.3 Study III

The RCT design is considered as the gold standard, and the only design that allows direct causal interpretation. Internal validity is therefore often warranted, despite the fact that more motivated and health-conscious people participate in these types of trials than those who are less motivated or health-conscious. External validity, also called generalisability, is, however, challenged by such self-selection into a RCT. Of the 1 007 nurses eligible to participate in the PRIMROSE trial, only 129 (12.8%) agreed to take part. A questionnaire among participating and non-participating nurses revealed some, albeit small, differences with regard to earlier MI education and specialist training. Among eligible families (n=1 867), more than 70% agreed to participate, which is considered a high proportion for this kind of intervention, where recruitment is often troublesome [102]. Of the 498 families declining to participate, 434 (87.1%) did answer a brief anonymous questionnaire, indicating only small differences with regard to maternal age, country of birth and highest education. Still, the PRIMROSE population may not be entirely representative of the general Swedish population. Among the PRIMROSE sample, around 10% of first-time mothers were born outside Sweden, in contrast to almost 20% in the general population. This difference is explained by the eligibility criteria of the trial. Because of the structure and content of the intervention, participants needed to be proficient in Swedish, which may not have been the case for recently migrated families.

Another limitation of the PRIMROSE intervention study concerns intervention delivery. As described earlier in detail, nurses received extensive education and practical training in MI prior and during the intervention period. Still, the vast majority did not reach beginning proficiency thresholds on any of the indicators of MI proficiency [103, 104]. This said, the proficiency thresholds used to evaluate MI skill levels are based on expert opinions only and have not been further validated. Also, there is no formal validation of the level of MI adherence needed to initiate a behaviour change among clients [105]. Yet, due to a possible lack of MI proficiency among nurses, we might have been unable to truly evaluate MI's efficacy.

The assessment of dietary intake and habits is known to be susceptible to biases, including recall bias and social desirability bias. Also, we found in our relative validation study that we were unable to capture all food intakes equally well. Parents (often mothers) filled in the FFQ for the questions, and we saw that children and mothers' intake were correlated, which may be due to the fact that children at the age of 4 have meals together with parents and therefore were reported to eat similar food. It may, however, also reflect measurement error due to mothers' being proxy reporters. If measurement error is systematic, e.g., resulting from social desirability bias, an overestimation of the intervention effect would be expected. Another limitation with regard to dietary assessment is a lack of data, as we only measure food intake consumed within the family. At follow-up, 95% of the children attended day care. However, it can be assumed that the access to and quality of food (in terms of composition) offered at day care centres will be the same for both the intervention and control group and should therefore only marginally have influenced our results. Furthermore, the focus of the intervention was on parental influences (especially mothers' behaviours) and not behaviour change among day-care centre personnel.

6.2.4 Study IV

The aim of the systematic review was not to pool estimates from individual studies to provide an overall estimate of cost-effectiveness, but rather to explore the scope, quality and variation of methodologies used within the field. If there had been the intention and capacity to pool the data, there would have been a need to be aware of publication bias. Considering the continuously increasing number of intervention studies within the field of early childhood obesity prevention, the small number of economic evaluation is somewhat surprising, but may reflect a lack of effectiveness and consequently lack of cost-effectiveness.

6.2.5 Study V

The main limitations of the economic evaluation of PRIMROSE are the lack of a significant intervention effect, a lack of HRQoL data for young children, limited costing data, and also too short a time horizon.

CUA are often preferred as these allow decision-makers to make comparisons and set priorities across different health outcomes by having a common effect measure, e.g., a QALY. In order to calculate a QALY there is a need to measure HRQoL using a preference-based instrument, like the EQ-5D, which was not included in the baseline or follow-up questionnaire. Therefore, we conducted instead a trial-based cost-effectiveness analysis with BMI unit as effect measure.

Another limitation is the restriction of the trial period, which means that effects and costs are limited to the observation period, i.e., up to 4 years of age. A longer time horizon is especially relevant to the evaluation of primary prevention interventions, where health effects and cost savings may only be visible in adulthood. Having too short a time horizon can potentially underestimate differences between the intervention and control groups. Additionally, we lacked information on the employment status and salaries of parents. Instead, we used region- and age-specific average salaries reported by Statistics Sweden to estimate the productivity losses due to participation in the trial.

To capture possible savings of the interventions, it would have been desirable to have had data on participants' health care utilisation. It can be argued that the majority of direct health-related cost savings will occur in adult age. However, a recently published Australian study indicates that, as early as in infancy, differences in health care utilisation can be detected between normal-weight and overweight children [106]. The health care costs of obese children were found to be 1.6 (95% CI: 1.12; 2.34) times those of normal weight children after adjustment for maternal characteristics and SES [106]. Failing to take these costs into account may lead to an under-estimation of cost-effectiveness. The PRIMROSE intervention addresses both parents (especially mothers) and their children. While keeping in mind the limitations in the assessment of dietary intake, there was an indication that mothers in the intervention group also had a slightly healthier food intake than those in the control group at 4-year follow-up. In economic evaluations, and in intervention research in general, one may further want to consider how to capture these or other spillover effects. Yet, so far, "current methods of cost-effectiveness analysis implicitly assume zero spill over among social ties" [107]. Younger siblings will most likely also benefit from this type of intervention. However, unfortunately, we did not collect information on whether families had further children and could therefore not explore this possibility.

6.3 THE NOVELTY OF PRIMROSE

Today, the prevention of early childhood obesity is well-established and a growing research field. In 2006, when the PRIMROSE intervention study was developed, it was among the first to evaluate intervention in infancy. At that time, the majority of intervention studies targeted children in a school setting, and only a few looked at children in kindergarten. The primary care setting was suggested as an intervention arena of interest, but had been largely underutilised.

PRIMROSE in turn was an intervention study closely embedded in a CHC setting with a design feasible to implement in public health practice. MI at that time was not yet evaluated as a way to convey intervention programs in the field of childhood obesity, either in a treatment or in a prevention setting. A well-established theory at the time was SCT for behavioural change. Yet, the majority of obesity interventions were not based on any theory, which was at that time believed to be a critical component of an intervention's success.

6.4 RELATION TO OTHER STUDIES

The latest Cochrane review on early childhood obesity prevention is by Waters and colleagues from 2011 [83]. In that review, the authors looked at 55 studies including children under the age of 18. Their main conclusion was that interventions related to physical activity and nutrition have the potential positively to influence children's BMI. However, the vast majority of studies had been conducted in school settings, and only eight looked at children under the age of five. Peirson and colleagues updated this review in 2015 with an addition of 62 studies [108]. In their meta-analysis they showed a small, yet significant overall effect (standardised mean difference=-0.07, 95% CI: -0.10; -0.03). Since the Cochrane review in 2011, interventions in early childhood have increased. Of the 20 interventions included for this age group, only one health promotion intervention, for Arab-Israeli, kindergarten children had a significant intervention effect [109]. Despite the increase in the number of interventions, the authors acknowledge a lack of studies for this age group. Until today, we are not aware of any other study that has examined early childhood obesity prevention in a health care setting, applying MI with the intensity and duration as the PRIMROSE intervention trial. Yet, there are a number of other related studies, which we will discuss below.

6.4.1 Intervention Studies in Early Childhood

A recent systematic literature review looked at interventions in the "first 1000 days" [110]. The authors identified 26 unique completed interventions, but none of them were found to be effective in preventing obesity or overweight with regard to prevalence. However, nine were effective with regard to small decreases in BMI, which may lead to the prevention of obesity if applied in larger populations. Of the 9 effective interventions, seven were based on individual- or family-level behaviour changes through counselling. However, the authors acknowledge that the results should be interpreted with caution given that only two [111, 112] of the included studies fulfilled all the quality criteria [110].

6.4.2 Interventions Embedded in a Primary Care Setting

At the time when PRIMROSE commenced in 2008, the primary care setting was considered a promising new and underutilised intervention arena [83]. Now in 2018, many more studies

have been conducted in that setting. Seburg and colleagues conducted a systematic review of primary care based obesity interventions that included both treatment and prevention interventions [113]. Of the 31 included studies, 8 had a significant effect on children's weight. However, 7 of them were treatment studies. The one preventive intervention study focused on sleeping and feeding practices in the first year of life, and was based on assisting parents to differentiate between feelings of hunger and other forms of distress, on assisting parents with the introduction of solid foods, and on how to overcome barriers with regard to the introduction of healthy food items. At follow-up, infants in the intervention group had lower weight-for-length percentiles than those in the control group [114].

6.4.3 Interventions Studies using MI

There is some evidence that MI is effective for treating obesity in adults [115]. But, in children, the evidence for a treatment effect of MI is mixed [116, 117]. The use of MI for prevention of childhood obesity is rare, but has been suggested as a promising approach [74]. For early childhood obesity interventions, MI is often directed towards parents only. We are aware of only one other study that applied MI as a counselling approach for the prevention of obesity in pre-school children age [118]. At follow-up at age 2-5 years the children in the "Healthy Habits, Happy Homes" intervention group had a (borderline) significantly lower BMI compared to the control group (-0.40, 95% CI: -0.79; 0.00). The intervention went on for about 6 months, and involved four MI sessions and four telephone calls by health educators [118]. Compared to the PRIMROSE intervention, the RCT had a more selective approach to prevention focusing on socially disadvantaged groups. Another Swedish intervention study used MI for the prevention of overweight and obesity among slightly older children (6 years of age) in a school setting, showing similar results to those of the PRIMROSE trial, i.e., no intervention effects on BMI and objectively physical activity behaviour (in boys), and small intervention effects on dietary behaviour [119].

6.4.4 Intervention Studies Targeting Parents as Agents

The components of the PRIMROSE intervention mainly addressed parents, based on the fact that they are role models in their children's emerging health behaviours. Parent involvement is almost always involved in early childhood obesity prevention. The meta-analysis of Yavuz and colleagues focused on early childhood obesity interventions (both prevention and treatment) that included parents and identified 50 studies [120]. The overall effect size was small, but statistically significant. However, only 20% of the interventions included presented significant (short-term) effect sizes and none of the studies maintained their effect at later follow-ups [120]. Furthermore, the majority of the studies showing intervention effects that concerned treatment rather than prevention. Therefore, the overall significant effect size might have been driven by the treatment effect.

6.5 THE ROLE OF OBESITY PREVENTION

In theory, primary prevention of obesity *should* be the most effective way of addressing the global obesity epidemic. Yet, over the last decades, researchers have not been successful. The results of the PRIMROSE intervention are largely in line with other similar primary prevention interventions, showing, if any, only small effect sizes. This raises questions about how to take

this important research field forward in the future. The main questions are at what age to start, who to target, and which intervention components to use.

6.5.1 When to Start?

The significance of early prevention has been widely acknowledged. As mentioned previously, behaviours contributing to the development of obesity are established early in life and become difficult to change once established. The gestational period has been identified as a critical or sensitive period for the development of offspring's obesity [121]. A large body of evidence has demonstrated that gestational weight gain plays an important role in offspring's birth weight [36], and weight trajectory in early childhood [122]. Therefore, the early childhood period may be too late to intervene, and there is a growing interest in interventions commencing already during early pregnancy to promote a healthy gestational weight gain [110]. Also, the preconception period has been suggested as a relevant period [123]. However, a Cochrane systematic review from 2015 that focused specifically on pre-pregnancy interventions for overweight or obese women did not identify even a single randomised trial [124]. An observational study from Canada showed that a 10% lower BMI is associated with a clinical meaningful risk reduction in a number of negative pregnancy outcomes [125]. However, so far, there has been no study that has investigated the effect of preconception weight reduction or overweight prevention on offspring's development of overweight, but studies are on the way [126].

6.5.2 Who to Target?

It has been suggested that the lack of evidence for prevention effects might be due to dilution, which may be due to targeting the general population [127]. The underlying assumption is that intervention effects are limited to those at highest risk and become diluted by targeting the general population. Another concern with interventions that target the general population is the possible widening of health inequalities, because low educated or poor families most in need of preventive measures may be least likely to take part in such interventions [128]. In addition, targeting the general population may give less health benefits for scare public resources than interventions targeting the most vulnerable groups. Therefore, targeting those most susceptible has been suggested to be potentially more effective than targeting the general population [18]. However, evidence from targeted intervention studies is sparse, and a concern with targeted intervention that has been raised is the potential stigmatisation of individuals [129][19]. Furthermore, Rose, many years ago, rightly pointed out that, once the underlying causes are known and controllable, "susceptibility ceases to matter" [130].

6.5.3 Which Components to Use?

A lack of effectiveness may also be explained by focusing on wrong or less sufficient intervention components. The main focus of the PRIMROSE intervention was on physical activity and dietary habits. For early childhood obesity, it has recently been suggested moving beyond dietary habits and physical activity, and addressing other behaviours and determinants, e.g., children's sleeping habits or stress in young families. Short sleep duration has been shown to be associated with an increased risk (RR: 1.45, 95% CI: 1.14; 1.85) of obesity among children [131], and other sleep patterns (i.e., routines) or late bedtimes may contribute to a higher risk of developing obesity among children [132]. Stress has also been suggested as an

important factor. Both stress in parents (or parents' perceived stress) [133] and stress in early life (i.e., a lack of self-regulations skills [134]) are associated with an increased risk of obesity in childhood.

However, it may not be enough to focus on individual obesity-related behaviours, as these are known to be shaped by upstream contextual social determinants, which operate beyond the individual level [135]. As early as in the 1970s, it was shown that obesity in children is strongly associated with parental socio-economic position [136]. Income, education and neighbourhood influence knowledge, motivation and the ability to engage in healthy behaviours [137]. Therefore, multi-sectoral interventions that focus on the root causes of obesity are necessary. These require a shift in responsibilities and a stronger political commitment in the fight against obesity.

6.5.4 Implementation Issues

One explanation for the absence of an effect of the PRIMROSE intervention might lie in difficulties related to implementation fidelity [138]. Bohman and colleagues showed that the nurses who participated in the intervention had not reached beginners' proficiency in MI at the beginning of the intervention, when nurses had not yet received all supervision units [103]. Persson and colleagues analysed the sample once more after they received the full MI training programme and found only a marginal improvement in MI proficiency [104]. Process evaluation can help to further identify facilitators or hindering factors on the pathway between intervention and effects, and should ideally combine qualitative and quantitative research [139].

6.6 ANALYSIS AND INTERPRETATION OF INTERVENTIONS

6.6.1 Intention-To-Treat Analysis versus Per-Protocol Analysis

The validity of the RCT design may be challenged by non-compliance, protocol deviations, or dropout. Therefore, the CONSORT criteria suggest that RCT should be evaluated according to the intention-to-treat principle (ITT) to deal with non-random attrition [140]. In the ITT analysis, each randomised participant is included in the final analysis and is in the treatment group to which he/she is initially assigned, regardless of what happened after randomisation (e.g., cross-over or dropout). Results from ITT analyses show a more modest effect size, and acknowledge that non-compliance and protocol deviation may also occur in practice. However, ITT analyses are often challenged by a lack of outcome data. In addition to the results of an ITT analysis, it is often of interest to provide results from a so-called per-protocol analysis. A per-protocol (PP) analysis only includes participants who followed the protocol, and results provide therefore an indication of treatment efficacy. The CONSORT guidelines recommend that both ITT and PP analyses should be reported [140]. Yet, both types of analyses have their limitations, especially in the setting of health promotion interventions [141]. There is currently no consensus about how ITT should be performed in the absence of complete outcome data, and strictly speaking "no analysis with missing outcome data can be described as ITT" [142]. In order to not omit any participants from an analysis, missing values need to be imputed. In longitudinal data analysis, the "last observation carried forward" is still a widely used approach despite severe drawbacks of the method [143]. Another approach to dealing with the loss of follow-up is multiple imputation, which aims to account for some uncertainty in the estimates

[144]. A variant of ITT analysis is the modified ITT, sometimes also referred to as complete-case analysis, in which only subjects with available outcome measurements [145] are included. The basic assumption in providing a valid estimate is the "missing at random" or "missing at completely random" principle. However, there is currently no consensus regarding terminology and application of either ITT or modified ITT. The same difficulties apply with regard to PP analyses. Public health interventions are often complex by nature, and pragmatic trials in particular have a less stringent protocol. To strengthen analysis and reporting, an analysis plan should already be specified in the trial protocol. Furthermore, the reporting of trial results should include basic descriptive (baseline) variables comparing those who completed the trial to those who dropped out or were lost to follow-up.

6.6.2 Defining Success in Obesity Prevention Interventions

As mentioned above, the marginal and non-significant results of the PRIMROSE intervention with regard to BMI are, unfortunately, no exception and largely in line with other research in the field. Responses are two-sided. Whereas some argue that there is a need to "stop ineffective intervention" [146], another question concerns the concept of "success" in obesity intervention [147]. Is it reasonable to assume a BMI difference large enough to be detected when targeting the general population? Furthermore, in primary prevention, i.e., targeting those who are not yet overweight, we do not anticipate a BMI reduction, instead we would like to prevent unhealthy weight gain. Still, even in primary prevention interventions, individual BMI change is often the outcome measure of interest and a reduction in BMI units is interpreted wrongly as success.

It is well acknowledged that BMI is not the gold standard for assessing body fatness. The gold standard for measuring body fatness is the double labelled water method [148]. This method is, however, fairly demanding and costly, and therefore not applicable for screening or larger studies. Alternatives are air displacement plethysmography, which has a higher reliability and validity than BMI, but is not as demanding as underwater measurement [149]. Still, for larger epidemiological studies, other "surrogate techniques", including BMI, waist circumference, skinfold thickness, bioelectric impedance and hip-to-waist ratio are used to measure body fatness. Despite its limitations, BMI is still the preferred measure for capturing overweight and obesity. Especially in early childhood, the other measures may be of limited applicability in the assessment of obesity [150].

In health promotion intervention studies "softer" end points are often used. Health behaviour (e.g., diet and physical activity) that are believed to be on the causal pathway act as end points for intervention studies. One underlying reason for this is that it takes time before a behavioural change becomes manifest in terms of BMI change, and also that a behavioural change is of benefit beyond what can be measured in BMI units. Still, behavioural change is still not necessarily needed in universal primary prevention, especially in early childhood where behaviours are to be learned and established. Therefore, a focus on broader behaviours with regard to physical activity and dietary behaviour may be a step forward.

6.6.3 Economic Evaluations

Due to the high prevalence of overweight and obesity, large amounts of resources are spent on treatment and prevention. Given the scarce resources faced by health-care systems, economic

evaluations are needed to assist decision-makers in prioritising and determining how and where to get the best value for money. Economic evaluations are well-established and fairly straightforward for pharmaceutical drugs and other therapeutic interventions. However, it becomes more complex for preventive public health interventions, where the vast majority of costs and benefits are in the far future and go beyond the health care sector.

As described in Study IV, and also discussed by Frew and colleagues [151] in relation to the amount of intervention studies in the field of childhood obesity, the number of economic evaluations in that field is small. None of the 51 studies that were analysed in the latest Cochrane review from 2011 looked at the cost-effectiveness of the interventions [83]. In recent years, there has been a slowly growing number of cost-effectiveness analyses and a growing interest in the field. In Study IV we identified six studies, four of which were based on an individual intervention.

To capture all possible costs and benefits, data from various sources (e.g., clinical, epidemiological, and economic) are needed. Trial-based economic evaluations, as in Study IV, give an underestimation and may be therefore of limited use. To account for future, preferably life-time, costs we need to model our analyses to account for variability and uncertainty [54]. However, results from modelling studies are only as good as the data input, and high-quality input data for the costs and benefits in the field of early childhood obesity prevention are sparse. Therefore, current model-based economic evaluations in the field of childhood obesity also have severe flaws. In addition, to extrapolate the benefits at the life-time horizon, data on effect maintenance are required from early childhood to adulthood. It is unreasonable to expect intervention studies to have follow-up measurement up to 15 years later, yet the assumptions that are currently often made (i.e., full effect maintenance) are not supported by research. Furthermore, more data are needed about direct and indirect costs that already occur in childhood, e.g., health care utilisation and educational attainment [151].

6.7 IMPLICATIONS OF FINDINGS AND FUTURE RESEARCH

The results of Study I suggest that even when obese individuals lose weight to reach normal weight, their HRQoL does not become the same as for those who have been normal weight all the time. While acknowledging the fact that the weight reduction may have been unintentional, there is also the possibility that there are independent effects of past overweight/obesity status on current HRQoL, which cannot be reversed through weight reduction. This again highlights the importance of primary prevention. Future research should focus on disentangling intentional and unintentional weight change, and the role of past overweight/obesity status on health outcomes. Further research is also needed to study the association in (young) children.

The relative validation of the FFQ in Study II showed mixed results, which suggests the use of caution when measuring and interpreting food intake. Future validation studies may consider using an objective measurement to strengthen the interpretation of findings. Future research should also aim to identify new approaches to measuring dietary intake, given that traditional methods are known to be flawed. However, as Beaton stated as early as in 1994: "There is not, and probably never will be, a method that can estimate dietary intake without error" [152]. Therefore, future research should also look at how we can use existing methods, but appropriately identify and correct for errors in data collection and analysis [152].

The results of Study III leave us with the question whether the intervention indeed did not work, or whether other factors (e.g., implementation or evaluation) could explain the absence of intervention effects. Therefore, future intervention studies should put more emphasis on process evaluation, including both quantitative and qualitative research. In addition, other (softer) outcome measures, e.g., feeding styles or children's enjoyment, may add to current understanding of the causal pathway.

As shown in Study IV and V, even small intervention effects may be worth the money. This, however, depends largely on the assumptions made and on the WTP of decision-makers. Currently, there are no thresholds for obesity-related outcomes, which may be an area for future research. In addition, or alternatively, more effort should be placed on the inclusion of QALY measures for children. Also, effect maintenance and spill-over effects (e.g., towards siblings) need to be further investigated. Furthermore, more research is needed on the direct (e.g., health care utilisation) and indirect (e.g., educational attainment) costs of obesity/overweight in childhood.

6.8 CONCLUSION

There was no evidence for the effectiveness or cost-effectiveness of the population-based PRIMROSE prevention trial conducted at Swedish CHCs. However, overweight and obesity significantly impact the HRQoL of individuals, and even if people with overweight reach normal weight, their HRQoL is worse than that of those with constant normal weight, which highlights the need for obesity prevention. Given that the obesity crisis persists, new approaches may be required, and further research needed, to disentangle the "failure" of intervention from the "failure" of implementation and the "failure" of evaluation.

7 ACKNOWLEDGEMENTS

The PRIMROSE trial started long before I began my PhD. Therefore, I would also like to thank all people contributing to the data collection and project management and all nurses and families participating in the PRIMROSE trial.

Thank you, **Finn Rasmussen** for offering me the possibility to become a PhD student and for being always available for questions and feedback. I especially appreciate and enjoyed our discussions in my final year and you sharing your seemingly endless knowledge of childhood obesity with me.

Thank you, **Berit Heitmann** for being my co-supervisor. You invited me twice to your fantastic research team in Copenhagen. Your feedback and our discussions always motivated me and I always felt well supported.

Thank you, **Jeroen de Munter**, for not only being a co-supervisor but also being a great colleague. I truly enjoyed our discussion of epidemiology, statistics and research in general and I am thankful for your support throughout the years.

Thank you, **Niklas Zethraeus** for being my co-supervisor and sharing your knowledge in the field of health economics with me. I knew I could always rely on your expertise when needed.

Thank you, **Irene Glinos** for being my mentor. You and **Hans Maarse** are probably (very likely!) the main reasons why I wanted to enter the research world and do a PhD.

Thank you, **Per Tynelius**, not only for all the work you did for PRIMROSE but also your endless patience and sharing your statistical expertise with me.

Thank you to all co-authors: **Ata Ghaderi**, **Benjamin Bohman**, **Maria Westin**, **Lena Hansson**, **Elina Scheers Andersson**, **Margaretha Magnusson**, **Christel Larsson**, **Elinor Sundblom**, **Mikaela Willmer**, **Margareta Blennow**, **Lars Forsberg**, **Sanna Wallin**, **Susanne Mayer**. A special thanks to **Diana Sonntag** for sharing your expertise in economic evaluation with me and inviting me for a research stay at your department.

Thank you, **Anita Berglund** and the **SINGS network** for establishing with the Graduate School of Register-based Research a stimulating and fun learning environment.

Thank you all current and past colleagues at the **Department of Public Health Sciences** for teaching me the culture of "fika" and lots of fun times at work.

Thank you, **Åsa Norman** and **Jan Sauermann** for reading my "Kappa" and providing valuable feedback and **Jonathan Kimber** for final proof reading and language editing.

And finally, thank you to all my informal supervisors outside academia.

8 REFERENCES

1. WHO. *Obesity and overweight Fact sheet N°311*. 2015.
2. World Health Organization. *Global Database on BMI Index- BMI classification*. 16/03/2018; Available from: http://apps.who.int/bmi/index.jsp?introPage=intro_3.html.
3. Cole, T.J., et al., *Establishing a standard definition for child overweight and obesity worldwide: international survey.* Bmj, 2000. **320**(7244): p. 1240.
4. Onis, M.d., et al., *Development of a WHO growth reference for school-aged children and adolescents.* Bulletin of the World health Organization, 2007. **85**: p. 660-667.
5. Monasta, L., et al., *Defining overweight and obesity in pre-school children: IOTF reference or WHO standard?* Obesity Reviews, 2011. **12**(4): p. 295-300.
6. Marild, S., et al., *Prevalence trends of obesity and overweight among 10-year-old children in western Sweden and relationship with parental body mass index.* Acta Paediatr, 2004. **93**(12): p. 1588-95.
7. Rokholm, B., J.L. Baker, and T.I. Sorensen, *The levelling off of the obesity epidemic since the year 1999--a review of evidence and perspectives.* Obes Rev, 2010. **11**(12): p. 835-46.
8. Sundblom, E., et al., *Childhood overweight and obesity prevalences levelling off in Stockholm but socioeconomic differences persist.* Int J Obes (Lond), 2008. **32**(10): p. 1525-30.
9. Lissner, L., et al., *Trends in overweight and obesity in Swedish schoolchildren 1999-2005: has the epidemic reached a plateau?* Obes Rev, 2010. **11**(8): p. 553-9.
10. Bygdell, M., et al., *The rise and the recent decline of childhood obesity in Swedish boys: the BEST cohort.* Int J Obes (Lond), 2017. **41**(5): p. 807-812.
11. Stockholm Läns Landsting. *Årsrapport: Barnhälsovården i Stockholms län 2016*. 2017; Available from: https://www.vardgivarguiden.se/globalassets/behandlingsstod/barnhalsovard/bhv-rapporter/bhv-arsrapport-2016.pdf?IsPdf=true.
12. WHO. *Global Strategy on Diet, Physical Activity and Health, Recommended levels of physical activity for children aged 5 - 17 years*. 2018; Available from: http://www.who.int/dietphysicalactivity/factsheet_young_people/en/.
13. Tremblay, M.S., et al., *Canadian physical activity guidelines for the early years (aged 0–4 years).* Applied Physiology, Nutrition, and Metabolism, 2012. **37**(2): p. 345-356.
14. Commonwealth of Australia: Department of Health and Ageing. *National Physical Activity Recommendations for Children 0-5 Years*. 2010; Available from: http://www.health.gov.au/internet/main/publishing.nsf/content/9D831D9E6713F92ACA257BF0001F5218/$File/PA%20Rec%200-5%20yo%20-%20Web%20printable%20version.pdf.
15. NHS, *Factsheet 2: Physical activity guidelines for early years*. 2015.
16. Barnes, J., et al., *Letter to the Editor: Standardized use of the terms" sedentary" and" sedentary behaviours".* Applied Physiology Nutrition and Metabolism-Physiologie Appliquee Nutrition Et Metabolisme, 2012. **37**(3): p. 540-542.

17. Tremblay, M.S., et al., *Sedentary Behavior Research Network (SBRN)–Terminology Consensus Project process and outcome.* International Journal of Behavioral Nutrition and Physical Activity, 2017. **14**(1): p. 75.

18. Owen, N., et al. *Sedentary behavior: emerging evidence for a new health risk.* in *Mayo Clinic Proceedings*. 2010. Mayo Foundation.

19. Cliff, D.P., et al., *Objectively measured sedentary behaviour and health and development in children and adolescents: systematic review and meta-analysis.* Obes Rev, 2016. **17**(4): p. 330-44.

20. Cliff, D.P., et al., *Objectively measured sedentary behavior, physical activity, and plasma lipids in overweight and obese children.* Obesity (Silver Spring), 2013. **21**(2): p. 382-5.

21. Saunders, T.J., J.-P. Chaput, and M.S. Tremblay, *Sedentary behaviour as an emerging risk factor for cardiometabolic diseases in children and youth.* Canadian journal of diabetes, 2014. **38**(1): p. 53-61.

22. Jones, R.A., et al., *Tracking physical activity and sedentary behavior in childhood: a systematic review.* Am J Prev Med, 2013. **44**(6): p. 651-8.

23. Koster, A., et al., *Association of sedentary time with mortality independent of moderate to vigorous physical activity.* PLoS One, 2012. **7**(6): p. e37696.

24. Pearson, N., et al., *Associations between sedentary behaviour and physical activity in children and adolescents: a meta-analysis.* Obesity reviews, 2014. **15**(8): p. 666-675.

25. Livingstone, M.B.E., P. Robson, and J. Wallace, *Issues in dietary intake assessment of children and adolescents.* British Journal of Nutrition, 2004. **92**(S2): p. S213-S222.

26. Malik, V.S., M.B. Schulze, and F.B. Hu, *Intake of sugar-sweetened beverages and weight gain: a systematic review–.* The American journal of clinical nutrition, 2006. **84**(2): p. 274-288.

27. Wang, X., et al., *Fruit and vegetable consumption and mortality from all causes, cardiovascular disease, and cancer: systematic review and dose-response meta-analysis of prospective cohort studies.* Bmj, 2014. **349**: p. g4490.

28. Ledoux, T., M. Hingle, and T. Baranowski, *Relationship of fruit and vegetable intake with adiposity: a systematic review.* Obesity Reviews, 2011. **12**(5).

29. Shloim, N., et al., *Parenting styles, feeding styles, feeding practices, and weight status in 4–12 year-old children: a systematic review of the literature.* Frontiers in Psychology, 2015. **6**: p. 1849.

30. Hansson, L.M., et al., *Associations Between Swedish Mothers' and 3- and 5-Year-Old Children's Food Intake.* J Nutr Educ Behav, 2016. **48**(8): p. 520-529.e1.

31. Kröller, K. and P. Warschburger, *Associations between maternal feeding style and food intake of children with a higher risk for overweight.* Appetite, 2008. **51**(1): p. 166-172.

32. Faith, M.S., et al., *Parent-child feeding strategies and their relationships to child eating and weight status.* Obesity, 2004. **12**(11): p. 1711-1722.

33. Monasta, L., et al., *Early-life determinants of overweight and obesity: a review of systematic reviews.* Obesity reviews, 2010. **11**(10): p. 695-708.

34. Dietz, W.H., *Critical periods in childhood for the development of obesity.* The American journal of clinical nutrition, 1994. **59**(5): p. 955-959.

35. Kuh, D., et al., *Life course epidemiology.* Journal of Epidemiology & Community Health, 2003. **57**(10): p. 778-783.

36. Yu, Z., et al., *Pre-pregnancy body mass index in relation to infant birth weight and offspring overweight/obesity: a systematic review and meta-analysis.* PloS one, 2013. **8**(4): p. e61627.

37. Schack-Nielsen, L., et al., *Gestational weight gain in relation to offspring body mass index and obesity from infancy through adulthood.* International journal of obesity, 2010. **34**(1): p. 67.

38. Ino, T., *Maternal smoking during pregnancy and offspring obesity: Meta-analysis.* Pediatrics International, 2010. **52**(1): p. 94-99.

39. Yu, Z., et al., *Birth weight and subsequent risk of obesity: a systematic review and meta-analysis.* Obesity Reviews, 2011. **12**(7): p. 525-542.

40. Baird, J., et al., *Being big or growing fast: systematic review of size and growth in infancy and later obesity.* Bmj, 2005. **331**(7522): p. 929.

41. Harder, T., et al., *Duration of breastfeeding and risk of overweight: a meta-analysis.* American journal of epidemiology, 2005. **162**(5): p. 397-403.

42. Moorcroft, K.E., J.L. Marshall, and F.M. McCormick, *Association between timing of introducing solid foods and obesity in infancy and childhood: a systematic review.* Maternal & child nutrition, 2011. **7**(1): p. 3-26.

43. Chen, X., M.A. Beydoun, and Y. Wang, *Is sleep duration associated with childhood obesity? A systematic review and meta-analysis.* Obesity, 2008. **16**(2): p. 265-274.

44. Hubert, H.B., et al., *Obesity as an independent risk factor for cardiovascular disease: a 26-year follow-up of participants in the Framingham Heart Study.* Circulation, 1983. **67**(5): p. 968-977.

45. Seidell, J.C., *Obesity, insulin resistance and diabetes—a worldwide epidemic.* British Journal of Nutrition, 2000. **83**(S1): p. S5-S8.

46. Grotle, M., et al., *Obesity and osteoarthritis in knee, hip and/or hand: an epidemiological study in the general population with 10 years follow-up.* BMC musculoskeletal disorders, 2008. **9**(1): p. 132.

47. Bianchini, F., R. Kaaks, and H. Vainio, *Overweight, obesity, and cancer risk.* The lancet oncology, 2002. **3**(9): p. 565-574.

48. Singh, A.S., et al., *Tracking of childhood overweight into adulthood: a systematic review of the literature.* Obesity reviews, 2008. **9**(5): p. 474-488.

49. Steinberger, J. and S.R. Daniels, *Obesity, Insulin Resistance, Diabetes, and Cardiovascular Risk in Children An American Heart Association Scientific Statement From the Atherosclerosis, Hypertension, and Obesity in the Young Committee (Council on Cardiovascular Disease in the Young) and the Diabetes Committee (Council on Nutrition, Physical Activity, and Metabolism).* Circulation, 2003. **107**(10): p. 1448-1453.

50. Sorof, J.M., et al., *Isolated systolic hypertension, obesity, and hyperkinetic hemodynamic states in children.* The Journal of pediatrics, 2002. **140**(6): p. 660-666.

51. Invitti, C., et al., *Prevalence and concomitants of glucose intolerance in European obese children and adolescents.* Diabetes care, 2003. **26**(1): p. 118-124.

52. Eisenberg, M.E., D. Neumark-Sztainer, and M. Story, *Associations of weight-based teasing and emotional well-being among adolescents.* Arch Pediatr Adolesc Med, 2003. **157**(8): p. 733-8.

53. Strauss, R.S. and H.A. Pollack, *Social marginalization of overweight children.* Arch Pediatr Adolesc Med, 2003. **157**(8): p. 746-52.

54. Jia, H. and E.I. Lubetkin, *The impact of obesity on health-related quality-of-life in the general adult US population.* Journal of public health, 2005. **27**(2): p. 156-164.

55. Doring, N., J. de Munter, and F. Rasmussen, *The associations between overweight, weight change and health related quality of life: Longitudinal data from the Stockholm Public Health Cohort 2002-2010.* Prev Med, 2015. **75**: p. 12-7.

56. Odegaard, K., et al., *The Swedish cost burden of overweight and obesity–evaluated with the PAR approach and a statistical modelling approach.* International Journal of Pediatric Obesity, 2008. **3**(S1): p. 51-57.

57. Neovius, K., et al., *Obesity status and sick leave: a systematic review.* Obesity reviews, 2009. **10**(1): p. 17-27.

58. Klarenbach, S., et al., *Population-based analysis of obesity and workforce participation.* Obesity, 2006. **14**(5): p. 920-927.

59. Paraponaris, A., B. Saliba, and B. Ventelou, *Obesity, weight status and employability: empirical evidence from a French national survey.* Econ Hum Biol, 2005. **3**(2): p. 241-58.

60. Neovius, K., et al., *Lifetime Productivity Losses Associated with Obesity Status in Early Adulthood.* Applied health economics and health policy, 2012. **10**(5): p. 309-317.

61. Hamilton, D., A. Dee, and I. Perry, *The lifetime costs of overweight and obesity in childhood and adolescence: a systematic review.* Obesity Reviews, 2017.

62. Sonntag, D., S. Ali, and F. De Bock, *Lifetime indirect cost of childhood overweight and obesity: A decision analytic model.* Obesity, 2016. **24**(1): p. 200-206.

63. Sonntag, D., et al., *Estimating the lifetime cost of childhood obesity in Germany: Results of a Markov Model.* Pediatric obesity, 2015. **10**(6): p. 416-422.

64. Olbers, T., et al., *Laparoscopic Roux-en-Y gastric bypass in adolescents with severe obesity (AMOS): a prospective, 5-year, Swedish nationwide study.* The Lancet Diabetes & Endocrinology, 2017. **5**(3): p. 174-183.

65. Gill, T.P., *Key issues in the prevention of obesity.* Br Med Bull, 1997. **53**(2): p. 359-88.

66. Sim, L.J., L. Parker, and S.K. Kumanyika, *Bridging the evidence gap in obesity prevention: a framework to inform decision making*. 2010: National Academies Press.

67. Skouteris, H., et al., *Parental influence and obesity prevention in pre-schoolers: a systematic review of interventions.* Obesity reviews, 2011. **12**(5): p. 315-328.

68. Anzman, S.L., B.Y. Rollins, and L.L. Birch, *Parental influence on children's early eating environments and obesity risk: implications for prevention.* International Journal of Obesity, 2010. **34**(7): p. 1116-1124.

69. Elfhag, K., P. Tynelius, and F. Rasmussen, *Family links of eating behaviour in normal weight and overweight children.* International Journal of Pediatric Obesity, 2010. **5**(6): p. 491-500.

70. Prestwich, A., et al., *Does theory influence the effectiveness of health behavior interventions? Meta-analysis.* Health Psychology, 2014. **33**(5): p. 465.

71. Miller, W.R. and S. Rollnick, *Motivational interviewing: Helping people change.* 2012: Guilford Press.

72. Miller, W.R. and G.S. Rose, *Toward a theory of motivational interviewing.* American Psychologist, 2009. **64**(6): p. 527.

73. Lundahl, B., et al., *Motivational interviewing in medical care settings: A systematic review and meta-analysis of randomized controlled trials.* Patient Education and Counseling, 2013. **93**(2): p. 157-168.

74. Resnicow, K., R. Davis, and S. Rollnick, *Motivational interviewing for pediatric obesity: conceptual issues and evidence review.* Journal of the American Dietetic Association, 2006. **106**(12): p. 2024-2033.

75. Lindhe Söderlund, L., *Motivational interviewing in theory and practice.* 2010, Linköping University Electronic Press.

76. Bandura, A., *Self-efficacy: The exercise of control*. 1997, New York: Freeman.

77. Burke, B.L., H. Arkowitz, and M. Menchola, *The efficacy of motivational interviewing: a meta-analysis of controlled clinical trials.* Journal of consulting and clinical psychology, 2003. **71**(5): p. 843.

78. Campbell, K., et al., *Maternal self-efficacy regarding children's eating and sedentary behaviours in the early years: Associations with children's food intake and sedentary behaviours.* Pediatric Obesity, 2010. **5**(6): p. 501-508.

79. Britton, M., *Statens beredning för medicinsk utvärdering.* Förebyggande åtgärder mot fetma: en systematisk litteraturöversikt. Stockholm: Statens beredning för medicinsk utvärdering (SBU), 2005.

80. Doak, C.M., et al., *The prevention of overweight and obesity in children and adolescents: a review of interventions and programmes.* Obes Rev, 2006. **7**(1): p. 111-36.

81. Hesketh, K.D. and K.J. Campbell, *Interventions to prevent obesity in 0-5 year olds: an updated systematic review of the literature.* Obesity (Silver Spring), 2010. **18 Suppl 1**: p. S27-35.

82. Summerbell, C.D., et al., *Interventions for preventing obesity in children.* Cochrane Database Syst Rev, 2005(3): p. CD001871.

83. Waters, E., et al., *Interventions for preventing obesity in children (review).* Cochrane collaboration, 2011(12): p. 1-212.

84. Larson, N., et al., *What role can child-care settings play in obesity prevention? A review of the evidence and call for research efforts.* Journal of the American Dietetic Association, 2011. **111**(9): p. 1343-1362.

85. Thorpe, K.E., et al., *A pragmatic–explanatory continuum indicator summary (PRECIS): a tool to help trial designers.* Journal of clinical epidemiology, 2009. **62**(5): p. 464-475.

86. Anell, A., A.H. Glenngard, and S.M. Merkur, *Sweden: Health system review.* Health systems in transition, 2012. **14**(5): p. 1-159.

87. Drummond, M.F. and A. McGuire, *Economic evaluation in health care: merging theory with practice*. 2001: Oxford University Press.

88. Huang, I.C., D.A. Revicki, and C.E. Schwartz, *Measuring pediatric patient-reported outcomes: good progress but a long way to go.* Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation, 2014. **23**(3): p. 747-750.

89. Riley, A.W., *Evidence that school-age children can self-report on their health.* Ambul Pediatr, 2004. **4**(4 Suppl): p. 371-6.

90. Cremeens, J., C. Eiser, and M. Blades, *Characteristics of health-related self-report measures for children aged three to eight years: a review of the literature.* Quality of Life Research, 2006. **15**(4): p. 739-754.

91. Griebsch, I., J. Coast, and J. Brown, *Quality-adjusted life-years lack quality in pediatric care: a critical review of published cost-utility studies in child health.* Pediatrics, 2005. **115**(5): p. e600-e614.

92. Carter, R., et al., *Assessing cost-effectiveness in obesity (ACE-obesity): an overview of the ACE approach, economic methods and cost results.* BMC Public Health, 2009. **9**(1): p. 419.

93. Ludvigsson, J.F., et al., *External review and validation of the Swedish national inpatient register.* BMC public health, 2011. **11**(1): p. 450.

94. Barlow, L., et al., *The completeness of the Swedish Cancer Register–a sample survey for year 1998.* Acta oncologica, 2009. **48**(1): p. 27-33.

95. Zou, G., *A modified poisson regression approach to prospective studies with binary data.* Am J Epidemiol, 2004. **159**(7): p. 702-6.

96. de Boer, M.R., et al., *Testing for baseline differences in randomized controlled trials: an unhealthy research behavior that is hard to eradicate.* Int J Behav Nutr Phys Act, 2015. **12**: p. 4.

97. Moher, D., et al., *CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials.* BMJ, 2010. **340**: p. c869.

98. Hinkle, D.E., W. Wiersma, and S.G. Jurs, *Applied statistics for the behavioral sciences*. Vol. 663. 2003: Houghton Mifflin College Division.

99. Gasser, T., D. Gervini, and L. Molinari, *Kernel estimation, shape-invariant modelling and structural analysis.* Cambridge Studies in Biological and Evolutionary Anthropology, 2004: p. 179-204.

100. Reviews, U.o.Y.C.f. and Dissemination, *Systematic reviews: CRD's guidance for undertaking reviews in health care*. 2009: University of York, Centre for Reviews & Dissemination.

101. Nawaz, H., et al., *Self-reported weight and height.* American journal of preventive medicine, 2001. **20**(4): p. 294-298.

102. Walters, S.J., et al., *Recruitment and retention of participants in randomised controlled trials: a review of trials funded and published by the United Kingdom Health Technology Assessment Programme.* BMJ open, 2017. **7**(3): p. e015276.

103. Bohman, B., et al., *An evaluation of training in motivational interviewing for nurses in child health services.* Behavioural and Cognitive Psychotherapy, 2013. **41**(3): p. 329-343.

104. Persson, J.E., et al., *Proficiency in motivational interviewing among nurses in child health services following workshop and supervision with systematic feedback.* PloS one, 2016. **11**(9): p. e0163624.

105. Apodaca, T.R. and R. Longabaugh, *Mechanisms of change in Motivational Interviewing: A review and preliminary evaluation of the evidence.* Addiction (Abingdon, England), 2009. **104**(5): p. 705-715.

106. Hayes, A., et al., *Early childhood obesity: Association with healthcare expenditure in Australia.* Obesity, 2016. **24**(8): p. 1752-1758.

107. Fletcher, J. and R. Marksteiner, *Causal spousal health spillover effects and implications for program evaluation.* American Economic Journal: Economic Policy, 2017. **9**(4): p. 144-66.

108. Peirson, L., et al., *Prevention of overweight and obesity in children and youth: a systematic review and meta-analysis.* CMAJ Open, 2015. **3**(1): p. E23-E33.

109. Nemet, D., et al., *Health promotion intervention in Arab-Israeli kindergarten children.* Journal of Pediatric Endocrinology and Metabolism, 2011. **24**(11-12): p. 1001-1007.

110. Blake-Lamb, T.L., et al., *Interventions for Childhood Obesity in the First 1,000 Days A Systematic Review.* American Journal of Preventive Medicine, 2016. **50**(6): p. 780-789.

111. Daniels, L., et al., *Evaluation of an intervention to promote protective infant feeding practices to prevent childhood obesity: outcomes of the NOURISH RCT at 14 months of age and 6 months post the first of two intervention modules.* International journal of obesity, 2012. **36**(10): p. 1292.

112. *Short- and long-term effects of feeding hydrolyzed protein infant formulas on growth at ≤6 y of age: results from the German Infant Nutritional Intervention Study.* The American Journal of Clinical Nutrition, 2009. **89**(6): p. 1846-1856.

113. Seburg, E.M., et al., *A Review of Primary Care-Based Childhood Obesity Prevention and Treatment Interventions.* Curr Obes Rep, 2015. **4**(2): p. 157-73.

114. Paul, I.M., et al., *Preventing obesity during infancy: a pilot study.* Obesity (Silver Spring), 2011. **19**(2): p. 353-61.

115. Armstrong, M., et al., *Motivational interviewing to improve weight loss in overweight and/or obese patients: a systematic review and meta-analysis of randomized controlled trials.* Obesity reviews, 2011. **12**(9): p. 709-723.

116. Broccoli, S., et al., *Motivational Interviewing to Treat Overweight Children: 24-Month Follow-Up of a Randomized Controlled Trial.* Pediatrics, 2016. **137**(1).

117. Wake, M., et al., *Outcomes and costs of primary care surveillance and intervention for overweight or obese children: the LEAP 2 randomised controlled trial.* Bmj, 2009. **339**: p. b3308.

118. Haines, J., et al., *Healthy habits, happy homes: Randomized trial to improve household routines for obesity prevention among preschool-aged children.* JAMA Pediatrics, 2013. **167**(11): p. 1072-1079.

119. Nyberg, G., et al., *Effectiveness of a universal parental support programme to promote healthy dietary habits and physical activity and to prevent overweight and obesity in 6-year-old children: the Healthy School Start Study, a cluster-randomised controlled trial.* PLoS One, 2015. **10**(2): p. e0116876.

120. Yavuz, H.M., et al., *Interventions aimed at reducing obesity in early childhood: a meta-analysis of programs that involve parents.* Journal of Child Psychology and Psychiatry, 2015. **56**(6): p. 677-692.

121. Hill, B., H. Skouteris, and M. Fuller-Tyszkiewicz, *Interventions designed to limit gestational weight gain: a systematic review of theory and meta-analysis of intervention components.* Obesity Reviews, 2013. **14**(6): p. 435-450.

122. Mamun, A. and M. Mannan, *Gestational weight gain in relation to offspring obesity over the life course: a systematic review and bias-adjusted meta-analysis.* Obesity Reviews, 2014. **15**(4): p. 338-347.

123. Stephenson, J., et al., *Before the beginning: nutrition and lifestyle in the preconception period and its importance for future health.* The Lancet.

124. Opray, N., et al., *Directed preconception health programs and interventions for improving pregnancy outcomes for women who are overweight or obese.* Cochrane Database Syst Rev, 2015(7): p. Cd010932.

125. Schummers, L., et al., *Risk of adverse pregnancy outcomes by prepregnancy body mass index: a population-based study to inform prepregnancy weight loss counseling.* Obstetrics and gynecology, 2015. **125**(1): p. 133.

126. LeBlanc, E.S., et al., *Prepare, a randomized trial to promote and evaluate weight loss among overweight and obese women planning pregnancy: Study design and rationale.* Contemporary clinical trials, 2016. **49**: p. 174-180.

127. Olsen, N.J., et al., *The Healthy Start project: a randomized, controlled intervention to prevent overweight among normal weight, preschool children at high risk of future overweight.* BMC Public Health, 2012. **12**(1): p. 1.

128. Gordon, D., *Inequalities in health: the evidence presented to the independent inquiry into inequalities in health, chaired by Sir Donald Acheson*. 1999: Policy Press.

129. Puhl, R.M. and C.A. Heuer, *The stigma of obesity: a review and update.* Obesity, 2009. **17**(5): p. 941-964.

130. Rose, G., *Sick individuals and sick populations.* International journal of epidemiology, 2001. **30**(3): p. 427-432.

131. Li, L., et al., *Sleep duration and obesity in children: A systematic review and meta-analysis of prospective cohort studies.* J Paediatr Child Health, 2017. **53**(4): p. 378-385.

132. Miller, A.L., J.C. Lumeng, and M.K. LeBourgeois, *Sleep patterns and obesity in childhood.* Current opinion in endocrinology, diabetes, and obesity, 2015. **22**(1): p. 41-47.

133. Parks, E.P., et al., *Influence of Stress in Parents on Child Obesity and Related Behaviors.* Pediatrics, 2012. **130**(5): p. e1096-e1104.

134. Graziano, P.A., S.D. Calkins, and S.P. Keane, *Toddler self-regulation skills predict risk for pediatric obesity.* Int J Obes (Lond), 2010. **34**(4): p. 633-41.

135. Lakerveld, J. and J. Mackenbach, *The upstream determinants of adult obesity.* Obesity facts, 2017. **10**(3): p. 216-222.

136. Stunkard, A., et al., *Influence of social class on obesity and thinness in children.* Jama, 1972. **221**(6): p. 579-584.

137. McLaren, L., *Socioeconomic status and obesity.* Epidemiologic reviews, 2007. **29**(1): p. 29-48.

138. Damschroder, L.J., et al., *Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science.* Implementation science, 2009. **4**(1): p. 50.

139. Moore, G., et al., *Process evaluation in complex public health intervention studies: the need for guidance.* Journal of Epidemiology and Community Health, 2014. **68**(2): p. 101-102.

140. Begg, C., et al., *Improving the quality of reporting of randomized controlled trials: the CONSORT statement.* Jama, 1996. **276**(8): p. 637-639.

141. Organization, W.H., *Health promotion evaluation: recommendations to policy-makers: report of the WHO European Working Group on Health Promotion Evaluation.* 1998.

142. White, I.R., J. Carpenter, and N.J. Horton, *Including all individuals is not enough: lessons for intention-to-treat analysis.* Clinical trials (London, England), 2012. **9**(4): p. 396-407.

143. Lachin, J.M., *Fallacies of Last Observation Carried Forward.* Clinical trials (London, England), 2016. **13**(2): p. 161-168.

144. Sterne, J.A.C., et al., *Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls.* BMJ, 2009. **338**.

145. White, I.R., et al., *Strategy for intention to treat analysis in randomised trials with missing outcome data.* Bmj, 2011. **342**: p. d40.

146. Anchondo, I.M., et al., *MA-CORD: Stop Ineffective Childhood Obesity Inverventions.* Childhood Obesity, 2015. **11**(6): p. 731-732.

147. Armstrong, S.C. and A.C. Skinner, *Defining "Success" in Childhood Obesity Interventions in Primary Care.* Pediatrics, 2016. **138**(4).

148. Lohman, T.G., *Assessment of body composition in children.* Pediatric Exercise Science, 1989. **1**(1): p. 19-30.

149. de Vargas Zanini, R., et al., *Body Fat in Children Measured by DXA, Air-Displacement Plethysmography, TBW and Multicomponent Models: A Systematic Review.* Maternal and child health journal, 2015. **19**(7): p. 1567-1573.

150. Krebs, N.F., et al., *Assessment of child and adolescent overweight and obesity.* Pediatrics, 2007. **120 Suppl 4**: p. S193-228.

151. Frew, E., *Economic Evaluation of Childhood Obesity Interventions: Reflections and Suggestions.* Pharmacoeconomics, 2016. **34**(8): p. 733-40.

152. Beaton, G.H., *Approaches to analysis of dietary data: relationship between planned analyses and choice of methodology.* The American journal of clinical nutrition, 1994. **59**(1): p. 253S-261S.